SOVIET ASTRONOMY

AJ

VOLUME 1, NUMBER 3 MAY-JUNE 1957

ON THE OBSERVATION OF ARTIFICIAL SATELLITES

A. A. Mikhailov

During the third International Geophysical Year, as is known, artificial satellites will be put into orbit around the Earth. The observation of the motion of these satellites will give valuable information on the upper layers of the atmosphere, the shape and constitution of the Earth.

A difficult and responsible task to organize the visual observations of the artificial satellites faces the astronomers of the Soviet Union.

The first satellites will have an apparent visual magnitude between 4 and 9 and a period of revolution of about one and a half hours. They will be observable only during the morning and evening twilight and because of their high angular velocity (up to 1° per second) the observation time at any point during transit will not be longer than two minutes.

Because of the rotation of the Earth and its departure from sphericity, the satellite orbit will be systematically displaced relative to the Earth's surface.

It is highly improbable that at one and the same place, during the same twilight, two consecutive transits could be observed.

For the visual observation of the satellites it is proposed to organize a network of special stations with 10-20 observers, not necessarily professional astronomers, each of whom will look at a small area of the sky by means of a standard telescope having an entrance pupil of 50 mm and a field of view of about 11°. The time of transit of a satellite across a definite point on the celestial sphere will be recorded by some kind of a registering device coupled to the time signals.

It is possible that some of the satellites will be equipped with pulsed gas-discharge lamps which will periodically give a light flash.

This will enable the visual observations to be carried on all night.

The Astronomical Council of the USSR Academy of Sciences appeals to all astronomical institutions, all the astronomers of the Soviet Union and members of the All-Union Astronomical and Geodetical Society with the request to take an active part in the preparations for the visual observation of artificial satellites.

Instructions and the special equipment for observation can be obtained through the Astronomical Council.

President, Astronomical Council AN SSSR,
Corresponding Member AN SSSR

ENERGY DISSIPATION, HEATING AND IONIZATION IN INTERSTELLAR GAS DUE TO SHOCK WAVES

S. B. Pikel'ner

The energy dissipation of hydromagnetic perpendicular shock waves is calculated assuming that the magnetic energy equals the kinetic energy of the wave and is considerably greater than the thermal energy of the gas. The role of the cosmic-ray pressure equal to the pressure of the field is considered. The presence of a field decreases the dissipation 20-30 times, the influence of the cosmic rays is small. If the energy of the field is larger than the kinetic energy, for instance, because of the influence of differential rotation, the dissipation decreases according to H^{-4}. The fraction of the thermal energy which goes into the ionization of hydrogen, has been calculated. One wave, in a field $H \approx 4 \cdot 10^{-6}$, with a concentration $n \approx 10^{-2}$ cm^{-3}, increases the temperature by 30,000° and ionizes 1-2% of hydrogen. The passage of consecutive waves can increase the ionization up to 20-50%; the mean temperature is about 10,000°. The degree of ionization strongly depends on the assumed physical conditions, the temperature depends much less on them.

In a previous paper [1] the author gave reasons for considering that the rarefied galactic gas between gas clouds is in a state of rapid motion and forms, not a planar, but a spherical subsystem. This picture is confirmed by the spherical distribution of the sources of nonthermal galactic radio emission. However, it encounters the difficulty pointed out by L. Spitzer [2], that a supersonic velocities the kinetic energy of the gas is rapidly converted into heat and the motions must be damped in a time of the order of the characteristic time of the motion. The problem of the spherical gas subsystem was investigated in detail in Reference [3] on the basis of the recent data on the radio emission from the galaxy. It was shown that the magnetic-field intensity in the spherical system is approximately equal to $6 \cdot 10^{-6}$ oersted in the lower layers and $3 \cdot 10^{-6}$ oersted in the upper layers, the mean concentration of the gas is about 10^{-2} cm^{-3} and slowly decreases towards the upper layers. It was also pointed out that the presence of the magnetic field, whose energy is comparable to the kinetic energy of the gas, must decrease the dissipation since in this case the velocity of sound will be of the order of the velocity of the gas motion. In the present work a quantitative estimate is made of the energy dissipation and the heating and ionization of the gas due to the heat evolved in hydromagnetic shock waves are evaluated.

For simplicity we will only consider perpendicular waves, i. e., waves in which the lines of force are parallel to the wave front. By a suitable choice of the reference system the gas velocity can be made perpendicular to the front. In this case the velocity of sound is equal to [4]

$$c_1 = \sqrt{\frac{dp}{d\rho} + V^2}, \text{ where } V = \frac{H}{\sqrt{4\pi\rho}}, \qquad (1)$$

where V is the velocity of the hydromagnetic waves. If $V^2 >> dp/d\rho$, then $c_1 \approx V$. At the same time, if the magnetic-energy density is on the average equal to the kinetic-energy density, then $V \approx v$ is the velocity of the gas. Consequently, the motions are at the speed of sound and the shock waves are weak.

The entropy change ΔS in a weak hydromagnetic shock wave was investigated in Reference [4].

If $x = \rho^{-1}$ and T_1 is the temperature before compression, then

$$\Delta S = -\frac{(\Delta x)^3}{12\,T_1}\left(\frac{\partial^2 p^*}{\partial x^2}\right)_S, \tag{2}$$

where

$$p^* = p + \frac{H^2}{8\pi}; \quad \Delta x = x_2 - x_1 \ll x_1.$$

Equation (2) is similar to the equation for the usual weak wave [5]. It is of interest to include into consideration the cosmic-ray pressure p_{cr}, which under the conditions prevailing in the galaxy is close to the pressure of the field. For this, we write down the usual conditions for the equality of the momentum, mass and energy flow on either side of the front [4], including in them p_{cr} and the energy density of the cosmic rays, equal to $3p_{cr}$:

$$p_1 + \rho_1 v_1^2 + \frac{H_1^2}{8\pi} + p_{1\,cr} = p_2 + \rho_2 v_2^2 + \frac{H_2^2}{8\pi} + p_{cr} \tag{3}$$

$$\rho_1 v_1 = \rho_2 v_2; \tag{4}$$

$$\varepsilon_1 \rho_1 v_1 + \frac{1}{2}\rho_1 v_1^3 + p_1 v_1 + \frac{v_1 H_1^2}{4\pi} + 3v_1 p_{1\,cr} =$$

$$= \varepsilon_2 \rho_2 v_2 + \frac{1}{2}\rho_2 v_2^3 + p_2 v_2 + \frac{v_2 H_2^2}{4\pi} + 3v_2 p_{2\,cr} \tag{5}$$

Here, v_1 and v_2 are the gas velocities relative to the front. In the coordinate system in which the gas before the front is stationary, $v_1 - v_2 = v$, the gas velocity behind the front in the same system; ϵ is the thermal energy of the gas. By the substitution

$$p^* = p + \frac{H^2}{8\pi} + p_{cr}; \quad \varepsilon^* = \varepsilon + \frac{H^2}{8\pi\rho} + \frac{2p_{cr}}{\rho} \tag{6}$$

the equations are transformed into the usual form. We introduce the variable $x = \rho^{-1}$.

It is necessary to establish the dependence of p_{cr} on $\underline{x}$ for what follows. Since **H** is parallel to the front, the particles, moving in spirals around the force lines, cannot quickly leave the region being compressed and we will assume that the particle concentration is proportional to the gas density. The field intensity in this case is also proportional to the density. According to the adiabatic invariant, an increase in H increases the component, perpendicular to H, of the energy of a relativistic particle proportionately to $H^{1/2}$. But in the wave being considered, it is this component of the pressure which is acting. Therefore we can assume that $p_{cr} \infty \rho H^{1/2} \infty \rho^{3/2} \infty$ $\infty x^{-3/2}$. Substituting into (6), we find after differentiation

$$\left(\frac{\partial \varepsilon^*}{\partial x}\right)_S = -p^*, \quad \left(\frac{\partial \varepsilon^*}{\partial S}\right)_x = \left(\frac{\partial \varepsilon}{\partial S}\right)_x = T. \tag{7}$$

After this, the derivation reduces to the one carried out in Reference [4] but with the appropriate value of p^*. From Equations (3) - (5), using the usual method, the velocity of sound can be obtained with p_{cr} taken into account. To do this we take the differences between the right- and left-hand sides of (3), (4) and (5), assuming these differences to be small, and divide by dv. Using the dependence of H and p_{cr} on ρ and consequently on $\underline{v}$, we find for $v_{sd} = c_1$ the expression

$$c_1 = \sqrt{\frac{dp}{d\rho} + \frac{H^2}{4\pi\rho} + \frac{3}{2}\frac{p_{cr}}{\rho}} \approx 1.3\,v, \tag{8}$$

if

$$p_{cr} \approx \frac{H^2}{8\pi} \approx \frac{1}{2}\rho v^2. \tag{9}$$

The inclusion of cosmic rays increases the velocity of sound still further and makes the motion subsonic.

The irreversible energy dissipation is

$$\Delta Q = T_1\,\Delta S = -\frac{1}{12}\left(\frac{\partial^2 p^*}{\partial x^2}\right)_S (\Delta x)^3 \text{ ergs/g} \cdot \text{sec}. \tag{10}$$

Differentiating (6), taking into account the dependence of H and p_{cr} on $\underline{x}$ and substituting into (9) we find

$$\Delta Q = -\frac{1}{12}\left[\gamma(\gamma+1)\frac{p_1}{\rho_1} + 6\frac{H_1^2}{8\pi\rho_1} + \frac{15}{4}\frac{p_{cr}}{\rho_1}\right]\left(\frac{\Delta x}{x_1}\right)^3 \approx -\frac{5}{6}\frac{H_1^2}{8\pi\rho_1}\left(\frac{\Delta x}{x_1}\right)^3. \tag{11}$$

Without the cosmic rays

$$\Delta Q \approx -\frac{1}{2}\frac{H_1^2}{8\pi\rho_1}\left(\frac{\Delta x}{x_1}\right)^3. \tag{12}$$

To calculate the dissipation we must find the magnitude of Δx. From (3) and (4) we find the difference between the pressures on either side of the front as

$$\begin{aligned} p_2 - p_1 + \frac{H_2^2}{8\pi} - \frac{H_1^2}{8\pi} + p_{2\,cr} - p_{1\,cr} &= \rho_1 v_1^2 - \rho_2 v_2^2 = \\ &= \rho_1 v_1 (v_1 - v_2) = \frac{\rho_1 v^2}{1 - \rho_1/\rho_2}. \end{aligned} \tag{13}$$

Neglecting gas pressure and introducing the quantity $\alpha = \frac{\rho_2}{\rho_1} = \frac{x_1}{x_2}$, then using (9) and the dependence of H and p_{cr} on ρ, we find

$$(\alpha^2 - 1 + \alpha^{3/2} - 1)\left(1 - \frac{1}{\alpha}\right) = 2, \tag{14}$$

from which $\alpha = 1.9$ and $\frac{\Delta x}{x_1} = \frac{1}{\alpha} - 1 = -0.47$. Without the cosmic rays, $\alpha = 2.19$ and $\frac{\Delta x}{x_1} = -0.54$. Substituting into (11) and (12) we find, respectively, that

$$\Delta Q = 0.044 \cdot 2 \frac{H_1^2}{8\pi\rho_1}, \quad \Delta Q = 0.040 \cdot 2 \frac{H_1^2}{8\pi\rho} \; (p_{cr} = 0). \qquad (15)$$

The dissipation is found to be almost the same in both cases. The quantity $2H_1^2/8\pi\rho_1$ is equal to the energy of the wave (magnetic + kinetic) per gram. Equations (15) show that the wave, having traversed a certain mass of gas, will lose about 4% of the energy associated with this mass due to reversible dissipation. Consequently, if the gas consists of masses in motion, whose characteristic time is equal to t_0, then the energy of the motion (with Condition (9)) will appreciably decrease in a time $t \approx 25\ t_0$. At the same time, in the absence of the field the damping time would have been of the order of t_0. Consequently, the presence of the field decreases the dissipation by a factor of 20-30. If the intensity of the field H_1 is for any reason higher than the value taken, for example, because of the differential galactic rotation which stretches the force lines, then, as can be easily shown, the dissipation rapidly diminishes as H_1^{-4}.

It is necessary to make a few qualifying remarks here. First of all, the interaction between shock waves, their intersection, can be important.* However, this apparently will not significantly change the results since, firstly, the regions of intersection occupy a relatively small volume and, what is more important, the field is increased in the shock wave, therefore, the second wave on encountering the high field only produces a small additional compression and correspondingly a small amount of heating. This, of course, does not refer to waves produced, for example, by the expansion of the dense envelopes from novae and supernovae. In this case the kinetic energy is considerably higher than the magnetic energy in front of the wave and the magnetic energy does not impede the compression. The second remark deals with the type of shock waves. We have investigated only the perpendicular waves, while in a real case the field is inclined to the wave front. The calculation of the dissipation in such waves will be carried out by the author in the future; for the time being it can be pointed out that the nomograms describing the properties of hydromagnetic shock waves [6] differ little, when the degree of compression is small, for angles between the field and wave normal greater than 45°.

Up to now only the irreversible dissipation of energy was considered. In addition to it, there is the reversible dissipation — the increase of the thermal energy through adiabatic compression and the increase in magnetic energy. Both of these energies are converted back into kinetic energy during the subsequent expansion of the compressed gas and, therefore, the total dissipation produced in this way does not increase even if each individual wave is retarded by the reversible dissipation. It is true that radiation emission can lead to an additional dissipation because the compressed gas will cool and will not expand. In our case, however, the adiabatic heating is small compared to the irreversible heating (see below). In addition, the gas pressure only plays a small role; the expansion will be caused by the magnetic pressure which, therefore, will once again be converted into kinetic energy.

Let us now consider the effect of the dissipation on the physical state of the gas. The irreversible increase in the temperature of the ionized and neutral gas as the result of the passage of one wave is equal to

$$\Delta T_i = \frac{2m_H}{3k} \Delta Q \approx \frac{2m_H}{3k} 0.1 \frac{H_1^2}{8\pi\rho} \approx \frac{0.1}{12\pi k} \frac{H_1^2}{n}. \qquad (16)$$

Assuming that $H_1 = 4 \cdot 10^{-6}$, $n = 10^{-2}$, we find that $\Delta T_i \approx 30{,}000°$. This value is strongly dependent on H_1, $\underline{n}$ and the relation between the kinetic and the magnetic energy of the gas. The temperature increase due to the adiabatic compression of a monatomic gas by a factor 2 ($\alpha \approx 2$) is equal to $0.6\ T_1$. It is true that the adiabatic heating also refers to electrons, so that in the case of a high degree of ionization this will double its relative influence. In general, it can be stated that if $T_1 < 20{,}000°$, the role played by adiabatic heating is small.

Through collisions the energy of the ions will be transferred to electrons and the latter lose energy by the excitation and ionization of hydrogen (the energy loss is $n_0 n_e \varphi(T_e)$ ergs/cm³ sec, where n_0 is the concentration of neutral hydrogen atoms) and by the excitation of metastable levels [OI], [OII], [NII], [SII] and others (the energy loss is $nn_e\psi(T_e)$ ergs/cm³ sec, where n is the concentration of all the hydrogen atoms). The function $\varphi(T_e)$ has been calculated in References [7] and [8] on the basis of cross sections calculated in the Born approximation and the function $\psi(T_e)$ has been calculated in Reference [9]. It does not depend strongly on the degree of ionization since in an HI region, at a temperature of 10,000°-15,000°, one electron in exciting [SII] loses an amount of energy which is comparable to the energy lost in exciting [OII] and [NII] in an HII region at the same temperature. The values of the functions $\varphi(T_e)$, $\psi(T_e)$ and $z_i\chi_1$, which characterize the energy going into the ionization of hydrogen, are given in the Table.

---

*This was pointed out by A. B. Severnyi during discussion of the present work.

T_e	10 000°	15 000°	20 000°	30 000°	50 000°
$\varphi(T_e)\cdot 10^{25}$	5.1	125	3200	$3.8\cdot 10^4$	$2.0\cdot 10^5$
$z_i\chi_1\cdot 10^{25}$	0.24	8	240	$3.9\cdot 10^3$	$4.3\cdot 10^4$
$\psi(T_e)\cdot 10^{25}$	78.2	155	272	—	—
$(T_i - T_e)\frac{n_e}{n}\cdot 10^2$	90°	$5.6\cdot 10^2$	$1.1\cdot 10^4$	$2.2\cdot 10^5$	$2.5\cdot 10^6$
$f(T_e)$	$2.8\cdot 10^{-3}$	$2.7\cdot 10^{-2}$	$6.7\cdot 10^{-2}$	0.10	0.21
$\frac{\Delta n_i}{n}(n_e \approx 10^{-2}n)$	0	$3.6\cdot 10^{-4}$	$1.5\cdot 10^{-3}$	$5.0\cdot 10^{-3}$	10^{-2}
$n_e(t_2 - t_1)$ in years ($n_e \approx 10^{-2}\ n$)	2650	1030	700	490	
$\frac{\Delta n_i}{n}(T_e = T_i)$	0	$3.8\cdot 10^{-4}$	$1.9\cdot 10^{-3}$	$0.7\cdot 10^{-2}$	$2.8\cdot 10^{-2}$
$n_e(t_2 - t_1)(T_e = T_i)$	2550	600	90	20	

Using the methods developed in Reference [9] we can estimate that in the time during which T_e reaches the value determined by the condition for the equilibrium of the energy transferred from the ions into excitation, using the assumed values of the parameters, the wave front will advance a distance of 0.01 ps in an HII region and ~ 1 ps in an HI region. Inasmuch as these distances are considerably smaller than the scale of the motions, we will consider that the equilibrium has been established.

The difference in the ion electron temperatures is approximately given by the expression

$$n_e k(T_i - T_e) = \left(\varphi + \frac{n}{n_0}\psi\right) n_e n_0 \tau_{ie}, \tag{17}$$

where τ_{ie} is the relaxation time for the difference in energy, given by [9];

$$\tau_{ic} \approx \frac{m_H}{2m}\,\frac{(kT)^{3/2} m^{1/2}}{\pi e^4 n_e L}\ ; \quad L = 2\ln\frac{kTD}{e^2}\ , \tag{18}$$

where $D = \sqrt{\dfrac{kT_e}{8\pi n_e e^2}}$ is the Debye radius. For $T \approx 10^4$ and $n_e \sim 10^{-4}$, $L \sim 45$ and $\tau_{ie} \approx 15T_e^{3/2}n_e^{-1}$. Substituting this into (17), we calculate $T_i - T_e$ as a function of T_e (see the Table, where n_e has been taken as $n_e = 10^{-2}\,n$). From the Table it can be seen that for $T_e \approx$ 10,000° - 15,000° the temperature difference is not large, while for $T_e >$ 20,000° the difference rapidly increases. It is true that at such a temperature the rate of ionization of hydrogen increases, which will lower somewhat the temperature difference.

Let us now investigate the action of the wave on the gas, which we will consider to be weakly ionized ($n_e \approx 0.01\,n$). This assumption simplifies the presentation of the calculations, but is unimportant for the final conclusions. As the result of the passage of the wave T_i will increase by ~ 30,000° and, according to the Table, T_e will be about 20,000°. The energy of the gas will be spend on the ionization of hydrogen and on the excitation of hydrogen and other atoms. The fraction expended in ionization will be

$$f(T_e) = \frac{Z_i\chi_1}{\varphi(T_e) + \frac{n}{n_0}\psi(T_e)}, \tag{19}$$

which is given in the table. For T = 10,000° it is very small. Taking into account the meaning of the quantity $f(T_e)$, we can write down the obvious equation

$$-nkd\,T_i f(T_e) = \chi_1\,dn_i, \tag{20}$$

from which the increase in the ionization as the result of the passage of the wave is

$$\frac{\Delta n_i}{n} = \frac{k}{\chi_1}\int\limits_{10000}^{T_0} f(T_e)\,dT, \tag{21}$$

where the difference $T_i - T_e$ must be taken into account. The values of $\Delta n_i/n$ for various T_0 (for which the numbers in the first row must be taken) are given in the table with $n_e \approx 10^{-2}n$. We see that the average wave increases the ionization by 0.5-1% and that the value taken for the ionization is not very important.

The change of the gas temperature behind the wave front can be written in the form

$$\frac{3}{2}\frac{dT_i}{dt} = -\frac{1}{k}\left(\frac{n_0}{n}\varphi + \psi\right)n_e, \tag{22}$$

from which the change in temperature from T_2 to T_1 takes place in the time

$$t_2 - t_1 = \frac{3}{2n_e}\int\limits_{T_1}^{T_2}\frac{kdT}{\frac{n_0}{n}\varphi(T_e) + \psi(T_e)}, \tag{23}$$

where, once again, it is necessary to take into account the value of $T_i - T_e$, which is taken for $n_e = 0.01\,n$. The time in years during which the gas cools from 15,000 to 10,000°, from 20,000 to 15,000°, etc., is given in the table. The gas cools rapidly to 20,000° and then more slowly to 15,000° and 10,000° and further. In a fairly large interval of the degree of ionization (if it exceeds 5-10%), in practice the gas temperature in the time between the two traversals by a wave will fall to 15,000-20,000°. Consequently, the subsequent wave will heat the gas to ~ 50,000° and ionize it by 1-1.5%, after which the temperature again falls to 15,000-20,000°. Each successive wave will ionize 1-2% of the gas. The increase of the ionization decreases the difference between the ion and electron temperatures; the magnitude of Δn_i increases somewhat, while $t_2 - t_1$ will decrease. To show the magnitude of this effect the limiting values of these quantities are calculated assuming $T_e = T_i$ (but $n_0 > n_e$) and given in the table. Each successive wave ionizes about 2% of the gas in this case.

The increase of the ionization will stop when the number of recombinations in the time between two traversals by a wave is equal to the number of ionizations produced by one wave. The value of the ionization can be estimated from the equation

$$x^2 n^2 C_H t_0 \approx 0.02\, n, \tag{24}$$

where $C_H(T_e) \approx 3.5 \cdot 10^{-13}$ is the probability for recombination and $t_0 \approx 3 \cdot 10^{13}$ sec is the characteristic time. From (24) $x \approx 0.4$, i. e., in this case the ionization reaches an equilibrium value at about 40 %. Unfortunately, the value obtained depends rather strongly on the parameter values taken. If the density or the field intensity are different from the values taken by several times, so that temperature increase produced by the wave is, for example, 10,000°, then the limiting value of the ionization is ~ 10 %. The value of t_0 affects the result to a lesser degree. The final temperature of the gas depends little on the parameters taken and in almost all of the cases it is equal to 10,000-15,000°. If however, the dissipation greatly increases or if the density is so small that the wave will raise the gas temperature by 80,000-100,000°, then the hydrogen will be highly ionized, the cooling will be less and the temperature will fall slowly.

If it is possible to estimate the gas temperature or the hydrogen ionization from observations, then the above calculations can be used for estimating the magnitude of the dissipation and increasing the accuracy of the values of $\underline{n}$ and H_1.

In the absence of the field, ΔQ and, accordingly, ΔT, will be higher by a factor 20-30. Even the first wave ionizes the gas by 30-40%, and the subsequent waves will produce an almost complete ionization.

Crimean Astrophysical Observatory
of the Academy of Sciences USSR

Received December 11, 1956

LITERATURE CITED

[1] S. B. Pikel'ner, Izvest. Krymsk. Astrofiz. Obs. 10, 74 (1953); Doklady Akad. Nauk SSSR 88, 229 (1953).

[2] L. Spitzer, Astrophys. J. 124, 20 (1956).

[3] S. B. Pikel'ner and I. S. Shklovskii, Astron. J. (USSR) 34, 145 (1957).*

[4] F. Hoffmann and E. Teller, Phys. Rev. 80, 692 (1950); Problems of Modern Physics [Russian translation] (IL, 1954) Vol. 2.

[5] L. D. Landau and E. M. Lifshits, The Mechanics of Continuous Media [in Russian] (Gostekhizdat, 1954) p. 399.

[6] H. L. Helfer, Astrophys. J. 117, 177 (1953); Problems of Modern Physics [Russian translation] (IL, 1954) Vol. 2.

[7] S. A. Kaplan and S. I. Gopasiuk, Tsirk. L'vovskoi Astr.Obs. 1953, No. 25.

[8] J. Chamberlain, Astrophys. J. 117, 387 (1953).

[9] S. B. Pikel'ner, Izvest. Krymsk. Astrofiz. Obs. 11, 93 (1954); Doklady Akad. Nauk SSSR 95, 1157 (1954).

---

*[Soviet Astronomy – AJ, p. 149]

SHOCK WAVES IN INTERSTELLAR SPACE

III. HYDROMAGNETIC DISCONTINUITIES

S. A. Kaplan

The formation of discontinuities in the interstellar magnetic field, associated with shock waves in the interstellar medium, is studied. The refraction of the magnetic lines of force and the variation of the magnetic field in an inclined hydromagnetic discontinuity is determined by Equations (9) - (11). The density discontinuity in a perpendicular magnetic shock wave (i. e., when $\mathbf{v}\perp\mathbf{H}$) is calculated [see Equations (18) - (20)]. The influence of the magnetic field on the ionization discontinuities is discussed briefly[(21) - (24)]. The automodel problem is formulated for the motion of hydromagnetic discontinuities.

At the present time it can be considered established that in interstellar space there exists a magnetic field whose energy density is comparable to the kinetic-energy density of the motions of the interstellar medium $(H^2 \approx 4\pi\rho v^2)$. On the other hand, since in interstellar space shock waves and rapid changes in the ionization are frequently produced, then the appearance of magnetic discontinuities, i. e., jumps in the values of the parameters characterizing the magnetic field, must also be expected.

The theory of hydromagnetic shock waves has been developed by a number of authors (Hoffman and Teller [1], Helfer [2], Syrovatskii [3], Lust [4], Kaplan [5]). In all of this work, with the exception of [5], it was assumed that in the passage of gas through a magnetic discontinuity the total energy flux is conserved. We have already seen, however, that because of radiation emission in interstellar space this condition often does not hold [6, 7]. In the present work we will investigate, in greater detail than in [5], the formation of magnetic discontinuities in the presence of radiation emission, i. e., within the framework of ideas developed in References [6, 7].

1. Inclined Magnetic Discontinuities

Hydromagnetic shock waves, in contrast to the usual hydrodynamic shock waves, have one important feature: here, because of the presence of the magnetic field, there can take place a simultaneous transfer of the tangential, as well as of the normal,component of the momentum flux during the passage of the gas through the discontinuity. Therefore, tangential hydromagnetic discontinuities (the transfer of only the tangential component of the momentum) are linked with perpendicular shock waves (the transfer of only the mormal component of the momentum) by means of inclined hydromagnetic shock waves (the transfer simultaneously of the tangential and normal components of the momentum) [1, 3].

It must be pointed out that inclined hydromagnetic shock waves can be considered as a generalization of inclined hydrodynamic shock waves. In the latter, however, by a transformation of the coordinate system it is always possible to go to the case of a perpendicular wave, since in inclined shock waves the tangential component of the velocity is the same on both sides of the discontinuity. The necessity for the investigation of the inclined shock waves only arises in the case of interactions between them or with obstacles. In the case of inclined hydromagnetic shock waves, in general, we cannot make such a transformation, since the tangential components of the velocity are not equal here, and we must investigate the inclined hydromagnetic waves as such even in free space.

In addition to shock waves, in hydromagnetic dynamics there can also exist another form of discontinuity at which there is a jump only of the direction of the magnetic-intensity and velocity vectors. In this case, the velocities on either side of the discontinuity are $v = H\sqrt{4\pi\rho}$ and, therefore, such a discontinuity is called a hydromagnetic wave. For a classification of hydromagnetic discontinuities see References [1-4].

To simplify the investigations, we will assume that the discontinuities are plane surfaces and we will restrict ourselves to the case of stationary hydromagnetic discontinuities.

In the general case the magnetic-field intensity vector **H**, and the velocity vector **v**, subtend the angles φ and φ', which are not equal, with the normal to the wave front. Transforming to a coordinate system moving relatively to the first with a velocity

$$V = \mathbf{v} - \frac{v\cos\varphi'}{H\cos\varphi}\cdot\boldsymbol{H}, \tag{1}$$

we obtain a description of the motion in which vectors of the velocity and of the magnetic-field intensity are parallel to each other, both before the passage of the wave front and after. The first condition follows immediately from (1) and the second from the condition for the conservation of $[\mathbf{v}H]$ in the passage through the discontinuity. In the following we will investigate the motion in this coordinate system and we will restrict ourselves to the planar problem, i. e., we assume that the vectors **v** and H, on both sides of the discontinuity, and the normal to the wave front lie in the same plane. This condition holds in all cases with the exception of hydromagnetic waves, which will not be investigated here. It must be pointed out that the transformation (1) is not possible when the vectors **v** and **A** are perpendicular to one another.

A separate investigation of this case, called a perpendicular shock wave, will be carried out in the next paragraph.

The method for investigating the discontinuities in the gas motions are well known: we choose two references on either side of the front in such a way that the region in which the parameters change sharply is situated between the surfaces (in the conditions of interstellar space this means that the layers of gas in which radiation emission takes place and the temperature does not correspond to the radiation field must be between the reference surfaces) and then we write down the conservation equation for the mass and momentum flux for these two surfaces. Because of the radiation emission the conservation of energy does not take place and is replaced by the assignment of the temperature, T, on both reference surfaces.

Let us denote by v_1 and v_2 the absolute values of the gas velocity in the coordinate system (1) before and after the passage of the discontinuity, by H_1 and H_2 the absolute values of the magnetic-field intensity and by φ_1 and φ_2 the angles between the direction of gas motion, or the magnetic-force lines, and the normal. Therefore, $\varphi_2 - \varphi_1$ is the angle of "refraction" of the flow and magnetic-force lines (independently of the choice of the coordinate system).

With these symbols, the equation expressing the conservation of the mass current will be in the form:

$$J = \rho_1 v_1 \cos\varphi_1 = \rho_2 v_2 \cos\varphi_2, \tag{2}$$

the equation for the conservation of the momentum current will be:

$$\rho_1\left(RT_1 + v_1^2\cos^2\varphi_1 + \frac{H_1^2}{8\pi\rho_1}\sin^2\varphi_1\right) = \rho_2\left(RT_2 + v_2^2\cos^2\varphi_2 + \frac{H_2^2}{8\pi\rho_2}\sin^2\varphi_2\right), \tag{3}$$

and the equation for the conservation of the tangential component of the total momentum current will be:

$$\rho_1 \sin \varphi_1 \cos \varphi_1 \left(v_1^2 - \frac{H_1^2}{8\pi\rho_1} \right) = \rho_2 \sin \varphi_2 \cos \varphi_2 \left(v_2^2 - \frac{H_2^2}{4\pi\rho_2} \right). \tag{4}$$

Here, T_1 and T_2 are the gas temperatures on either side of the discontinuity whose values must be given.

From the condition $\operatorname{div} \mathbf{H} = 0$ we obtain another equation:

$$H_1 \cos \varphi_1 = H_2 \cos \varphi_2. \tag{5}$$

In individual cases other additional conditions are imposed on the parameters of the discontinuity. Before we proceed to the investigation of these cases, let us transform Equations (2) - (5) into a more useful form:*

$$J^2 = \rho_2\rho_1 \left\{ \frac{\rho_2 RT_2 - \rho_1 RT_1}{\rho_2 - \rho_1} + \frac{H_1^2 \cos^2 \varphi_1}{8\pi(\rho_2 - \rho_1)} (\operatorname{tg}^2 \varphi_2 - \operatorname{tg}^2 \varphi_1) \right\}, \tag{6}$$

$$J^2 = \rho_2\rho_1 \frac{H_1^2 \cos^2 \varphi_1}{4\pi} \frac{\operatorname{tg} \varphi_2 - \operatorname{tg} \varphi_1}{\rho_1 \operatorname{tg} \varphi_2 - \rho_2 \operatorname{tg} \varphi_1}. \tag{7}$$

The investigation of these equations in their general form is unwieldy and, in addition, does not enable the peculiarities of the various types of discontinuities to be clearly seen. Therefore, we will investigate separately the simplest cases of the discontinuities which are of physical and astrophysical interest.

2. Magnetic Discontinuities

Let us at first assume that the gas pressure, both before and after the passage of the front of the discontinuity, is considerably less than the magnetic.

We will call such cases magnetic discontinuities. Then Equations (6) and (7) can be written in the form:

$$\frac{J^2}{\rho_1\rho_2} = \frac{H_1^2 \cos^2 \varphi_1}{8\pi(\rho_2 - \rho_1)} [\operatorname{tg}^2 \varphi_2 - \operatorname{tg}^2 \varphi_1] = \frac{H_1^2 \cos^2 \varphi_1}{4\pi} \frac{\operatorname{tg} \varphi_2 - \operatorname{tg} \varphi_1}{\rho_1 \operatorname{tg} \varphi_2 - \rho_2 \operatorname{tg} \varphi_1}. \tag{8}$$

To avoid any misunderstanding it must be pointed out that the use of (8) does not mean that inside the front of the discontinuity the gas pressure is also less than the magnetic; it can be greater there; it is only necessary that the conditions $p_1 \ll \frac{H_1^2}{8\pi}$ and $p_2 \ll \frac{H_2^2}{8\pi}$ hold.

It is not difficult to show that the system (8) has three solutions:

$$\begin{gathered} \operatorname{tg} \varphi_2 = \frac{1}{2}\left(\frac{\rho_2}{\rho_1} - 1\right) \operatorname{tg} \varphi_1 \pm \sqrt{\frac{1}{4}\left(\frac{\rho_2}{\rho_1} + 1\right)^2 \operatorname{tg}^2 \varphi_1 + \frac{\rho_2}{\rho_1} - 1}, \\ \operatorname{tg} \varphi_2 = \operatorname{tg} \varphi_1, \end{gathered} \tag{9}$$

which correspond to three different velocities of propagation of the disturbances in the magnetized medium.

The plus sign in the first solution (9) corresponds to an increasing magnetic field. In large magnetic discontinuities, in the case of not too small angles $\rho_2 \tan \varphi_1 >> \rho_1$, we have (restricting ourselves to the solution with increasing magnetic field):

---

* 'tg' ≡ 'tan' – Publisher.

$$\operatorname{tg} \varphi_2 \approx \frac{\rho_2}{\rho_1} \operatorname{tg} \varphi_1 \tag{10}$$

Combining this with Equation (5) we obtain:

$$\frac{H_2}{H_1} \approx \frac{\rho_2}{\rho_1} \sin \varphi_1 . \tag{11}$$

Thus, in strong magnetic discontinuities with radiation emission the flow and magnetic-force lines behind the front of the discontinuity are almost parallel to the front, and the magnetic-field intensity increases by a factor $\frac{\rho_2}{\rho_1}$ sin φ_1. For the other parameters of these waves we have, in the same approximation,

$$v_1^2 \approx \frac{H_1^2}{8\pi\rho_1} \frac{\rho_2}{\rho_1} \operatorname{tg}^2 \varphi_1; \quad v_2^2 = \frac{H_1^2}{8\pi\rho_1} \operatorname{tg}^2 \varphi_1 . \tag{12}$$

If $\varphi_1 \to 0$ the equations derived become meaningless. In this case a better approximation is the case of a perpendicular hydromagnetic shock wave which we shall consider below.

The equations obtained enable us to investigate and generalize, by the inclusion of a magnetic field, any problem arising in the theory of inclined shock waves, for example, the problem of the reflection of a magnetic discontinuity from a rigid nonconducting wall or the problem, equivalent to it, of the intersection of two magnetic discontinuities. Even if there are no fundamental difficulties involved here, the resulting equations are very unwieldy and their solution in explicit form is difficult. As an illustration, Figure 1 gives without derivation the solution of the problem of the intersection of two identical strong magnetic discontinuities at a small angle $\epsilon << 2\sqrt{\rho_1/\rho_2}$, where the jump in the density is $\rho_2 >> \rho_1$ and in the coordinate system used the line of the discontinuity is at rest.

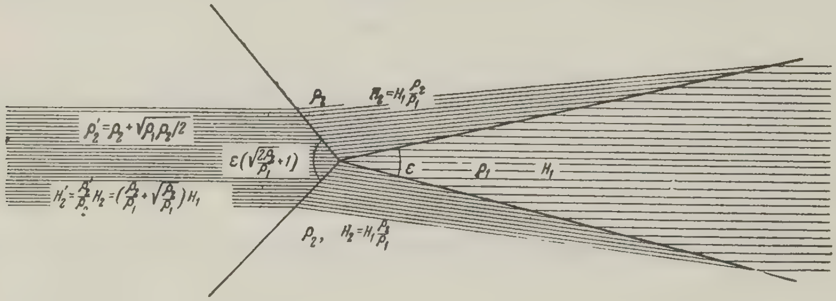

ρ_2
$H_2 = H_1 \frac{\rho_2}{\rho_1}$
$\rho_2' = \rho_2 + \sqrt{\rho_1 \rho_2 / 2}$
$\varepsilon(\sqrt{\frac{2\rho_2}{\rho_1}} + 1)$
ε
ρ_1
H_1
$H_2' = \frac{\rho_2'}{\rho_2} H_2 = (\frac{\rho_2}{\rho_1} + \sqrt{\frac{\rho_2}{\rho_1}}) H_1$
ρ_2, $H_2 = H_1 \frac{\rho_2}{\rho_1}$

Fig. 1

In concluding this section, we will make a few remarks on the very-strong shock waves in the general case. If $J \to \infty$ (with $v_1 \to \infty$), then it follows from (7) that

$$\frac{\rho_2}{\rho_1} \to \frac{\operatorname{tg} \varphi_2}{\operatorname{tg} \varphi_1}, \tag{13}$$

Combining (13) and (2), we find

$$v_1 \sin \varphi_1 \to v_2 \sin \varphi_2.$$

Thus, as the strength of the wave is increased, the tangential components of the velocities become comparable and the transfer of the tangential component of the momentum decreases. For very-strong inclined hydromagnetic shock waves, the usual theory of inclined shock waves can be a good approximation.

3. Perpendicular Hydromagnetic Shock Waves

If the magnetic field is parallel to the plane of the discontinuity, the transformation (1) becomes meaningless. In this case, called a perpendicular shock wave, it is advantageous to chose a coordinate system in which the front of the discontinuity is at rest and the velocity components are perpendicular to the front. Instead of (2) - (4), we now have

$$J = \rho_1 v_1 = \rho_2 v_2, \quad \rho_1\left(RT_1 + v_1^2 + \frac{H_1^2}{8\pi\rho_1}\right) = \rho_2\left(RT_2 + v_2^2 + \frac{H_2^2}{8\pi\rho_2}\right). \tag{14}$$

From the condition for the conservation of the quantity [**v H**] in going through the discontinuity, we obtain the equations:

$$v_1 H_1 = v_2 H_2 \quad \text{or} \quad \frac{H_1}{\rho_1} = \frac{H_2}{\rho_2} = b\sqrt{8\pi}, \tag{15}$$

where b is a constant (the "freezing-in" or "adhesion" integral [8]). Substituting (15) into (14), we obtain

$$\rho_1 (RT_1 + v_1^2 + b^2 \rho_1) = \rho_2 (RT_2 + v_2^2 + b^2 \rho_2). \tag{16}$$

Thus, the influence of the magnetic field on a perpendicular hydromagnetic shock wave can be taken into account by replacing the gas pressure ρRT by the total pressure $\rho RT + b^2\rho_2$. As was shown in Reference [8], an analogous method can be used in all problems for taking into account the influence of the magnetic field on the one-dimensional nonstationary gas flow in hydromagnetic dynamics.

We present, as an example, several equations which generalize the results obtained in References [6] and [7]:

a) The velocity of the hydromagnetic shock wave in a gas which is at rest before the passage of the wave is given by (see Equation (12) of Reference [6]):

$$v_1^2 = \frac{\rho_2}{\rho_1}\left[\frac{\rho_2 RT_2 - \rho_1 RT_1}{\rho_2 - \rho_1} + b^2(\rho_2 + \rho_1)\right]. \tag{17}$$

From this follows the equation determining the jump of the density in a hydromagnetic shock wave with radiation emission:

$$\frac{\rho_2}{\rho_1} = \frac{1}{2}\,\frac{v_1^2 + RT_1}{RT_2 + b^2\rho_1} \pm \sqrt{\frac{1}{4}\left(\frac{v_1^2 + RT_1}{RT_2 + b^2\rho_1}\right)^2 - \frac{v_1^2 + b^2\rho_1}{RT_2 + b^2\rho_2}}. \tag{18}$$

The difference between this equation and the analogous one, (13) of Reference [6], lies not only in the fact that here the magnetic field has been taken into account, but also in that in Equation (13) of Reference [6] v denotes the gas velocity behind the wave front, while in equation (18) v_1 denotes the velocity of the front itself. Since $v = v_1 - v_2 = v_1(1 - \rho_1/\rho_2)$, in the case of large jumps $\rho_2 \gg \rho_1$ and $v \approx v_1$. Here, in particular, it follows from (18) that

$$\frac{\rho_2}{\rho_1} = \frac{v^2}{RT_2 + H_1^2/8\pi\rho_1}. \tag{19}$$

If the initial magnetic field satisfies the condition $H_1^2 \gg 8\pi\rho_1 RT_2$, then

$$\frac{\rho_2}{\rho_1} \approx \frac{8\pi\rho_1 v_1^2}{H_1^2}, \tag{20}$$

or, numerically, with $H_1 \approx 10^{-5}$ gauss and $\rho_1 \approx 3 \cdot 10^{-24}$ g/cm^3,

$$\frac{\rho_2}{\rho_1} \approx \left(\frac{v}{11.5\,\mathrm{km/sec}}\right)^2. \tag{20'}$$

Thus, significant density jumps in the presence of a magnetic field can only take place along the direction of the magnetic field (in which case we can use, as before, Equation (15) of Reference [6] instead of Equation (20)), or with very-high velocities of gas ($v > 30$ km/sec), or with a high initial density ($\rho_1 > 3 \cdot 10^{-23}$ g/cm^3). Consequently, in the conditions of the interstellar space the magnetic field considerably lowers the compressibility of the gas and, apparently, has a significant influence on the formation of interstellar gas clouds.

b) A hydromagnetic ionization jump with radiation emission. By analogy to Equation (5) of Reference [7] we write ($\gamma = 1$, also neglecting the radiation pressure):

$$\begin{gathered} \rho_1 v_1 = \rho_2 \left[2RT_2 + 2b^2\rho\right]^{1/2} = m_\mathrm{H} C_\mathrm{H} M, \\ \rho_1 (v_1^2 + RT_1 + b^2\rho_1) = 4\rho_2 (RT_2 + b^2\rho_2), \end{gathered} \tag{21}$$

since in our case the velocity of propagation of the hydromagnetic disturbances is $v_2 = c_2 = [2RT_2 + 2b^2\rho_2]^{1/2}$. Solving (21), we obtain:

$$v_1 = [2(RT_2 + b^2\rho_2)]^{1/2} \pm \sqrt{R(2T_2 - T_1) + b^2(2\rho_2 - \rho_1)}, \tag{22}$$

where the density ρ_2 is determined by the following equation, also resulting from (21):

$$\rho_2^2 [2(RT_2 + b^2\rho_2)] = (m_\mathrm{H} C_\mathrm{H} M)^2, \tag{23}$$

after which all of the other parameters of the ionization discontinuity can be found. It must be pointed out, however, that the resulting equations are very unwieldy and, therefore, we will not consider the details of their solution. In the actual conditions of interstellar space, usually $RT_2 \approx b^2\rho_2$. However, if the energy density of the magnetic field is sufficiently high, so that $b^2\rho_2 \gg RT_2$, then (22) - (23) give

$$\rho_2 = (m_\mathrm{H} C_\mathrm{H} M / b^2)^{2/3}, \quad v_1 = (m_\mathrm{H} C_\mathrm{H} M b)^{2/3}. \tag{24}$$

Thus in this case all the parameters of the ionization discontinuity are determined, not by the gas temperature or the ionization energy, but by the intensity of the magnetic field. The investigation of such discontinuities can be easily carried out by the methods of Reference [7] and we will not consider them here in greater detail.

4. The Automodel Problem of the Distribution of Hydromagnetic Discontinuities

In hydromagnetic motions, first of all, the following two specifying parameters with independent dimensions must be considered: the density $[\rho_0]$ = g/cm^3 and the integral $[b]$ = cm$^{5/2}$/g$^{1/2}$ sec. Obviously, the hydromagnetic motion will only be an automodel one when all the other specifying parameters are combinations of these two. For example, the isothermal hydromagnetic motions are automodel since $[RT_0] = \text{cm}^2/\text{sec}^2 = [b^2\rho_0]$. The problem of the motion of an ionization discontinuity, investigated in detail in Reference [7], preserves its automodel character in the presence of a magnetic field. In this case, instead of Equation (15) of Reference [7], we have the system

$$\begin{aligned}
[\eta - V(\eta)]\frac{dV(\eta)}{d\eta} &= \frac{T_1}{T_2}\frac{d\ln R(\eta)}{d\eta} + \frac{b^2\rho_2}{R_{\text{gas}}T_2}\frac{dR(\eta)}{d\eta}, \\
[\eta - V(\eta)] - \frac{d\ln R(\eta)}{d\eta} &= \frac{dV(\eta)}{d\eta} + (N-1)\frac{V(\eta)}{\eta},
\end{aligned} \tag{25}$$

with the same notation as in Reference [7] (here R_{gas} is the gas constant). This system has the same solution, R = const, V = const for N = 1, as Equations (15) of Reference [7]. The boundary conditions, however, are more complicated here than in Reference [7]. Nevertheless, the solution of the problem of the propagation of an ionization discontinuity in the presence of a magnetic field does not present any fundamental difficulties.

5. Conclusion

It must be pointed out once more that in magnetic discontinuities with radiation emission a very large increase of the magnetic-field intensity, as well as of the density, can take place. This increase is associated with the effect of the "freezing-in" or of the "adhesion" of the magnetic field to matter, inasmuch as with an infinite conductivity (and this condition is assumed here) the magnetic-force lines are, so to say, frozen into matter, and when the gas density increases, the "density" of the magnetic-force lines also increases, i. e., the magnetic-energy density increases. If the gas particles do not move along the magnetic-force lines, the increase of the magnetic-field intensity is proportional to the density increase; in the presence of such a motion along the field, Equation (11) holds, where φ_1 is the angle between the magnetic field and the normal to the wave front.

In the present work, as also in References [6, 7], the question of the stability of magnetic discontinuities was not considered at all. It is obvious that not all of the conditions of flow investigated are stable. The general problem of the stability of discontinuities has, as yet, been insufficiently studied and we hope to deal with it in future papers.

L'vov Astronomical Observatory

Received May 2, 1956

LITERATURE CITED

[1] F. Hoffmann and E. Teller, Phys. Rev. 80, 697 (1950).

[2] H. L. Helfer, Astrophys. J. 117, 177 (1953).

[3] S. I. Syrovatskii, J. Exptl.-Theoret. Phys. (USSR) 24, 622 (1953).

[4] R. Lust, Z. Naturforsch 10a, 125 (1954).

[5] S. A. Kaplan, Astron. J. (USSR) 31, 358 (1954).

[6] S. A. Kaplan, Astron. J. (USSR) 33, 646 (1956).

[7] S. A. Kaplan, Astron. J. (USSR) 34, 183 (1957).*

[8] S. A. Kaplan and K. P. Staniukovich, Doklady Akad. Nauk. SSSR 95, 769 (1954).

---

*[Soviet Astronomy — AJ, p.183].

A SPECTROSCOPIC INVESTIGATION OF THE DEUTERIUM D_α LINE IN ACTIVE REGIONS ON THE SUN

A. B. Severnyi

The behavior of the λ6561.105 line in 28 short-lived grains of emission ("moustaches") and in the nearby undisturbed atmosphere was investigated with the Solar Tower Telescope at high resolving power (300,000) and dispersion (4 mm = 1 A). In the spectra of these grains, an average increase of the equivalent width of this line by 30% and the contrast by 1.7 % was found. It is shown that this effect cannot be attributed to accidental or systematic errors, but is connected with the strengthening of some solar line, which is heavily blended by the water vapor line λ 6561.105.

The appearance in sunspot spectra of a line in the same place as D_α is demonstrated in several cases of exceptionally low atmospheric water-vapor content, when λ 6561.105 was practically absent on the undisturbed disk. The wavelength of this unknown line 6561.09 A is precisely that of the red component of the fine structure of D_α. The equivalent width – 6.4 mA (if the line is attributed to D_α) indicates an increase of the $\frac{D}{H}$ relative abundance up to 10^{-4}. It is possible that there are sources of neutrons in active regions on the Sun.

In a recent paper [1] we presented spectroscopic evidence on the presence of a weak D_α-line of the heavy isotope of hydrogen in the solar spectrum. The presence of this line shows up as a small variation in the intensity and the profile of the line λ 6561.105 (attributed to the water vapor in the earth's atmosphere and masking the possible D_α-line) from the center to the eastern and western edges of the solar disc. From measurements of this effect we deduced the equivalent width of the possible D_α-line to be about 1.5 mA and the relative concentration of deuterium in the undisturbed solar atmosphere to be $D/H \approx 3\text{-}5 \cdot 10^{-5}$.

However, during investigations of the spectra of the short-lived continuous and line emission (the phenomenon of "moustaches") discovered by us [2] near the H_α-line, we have noticed that the line λ 6561.105 in regions of this emission differs appreciably from its continuation in the undisturbed atmosphere [3]. Moreover, it was found that in the sunspot spectrum it is also different from its continuation into the region next to the spot.

The profiles of the line λ 6561.105 were studied in the region of emission near H_α and in the neighboring regions of the undisturbed atmosphere, in the case of 28 active formations – mainly granules with emission "moustaches" and continuous emission, flares, and granules with continuous emission in the region of sunspot groups. Figure 1 is an example of one of the investigated strips showing the emission of "moustaches" and Table 1 gives all the data on the spectra investigated (Columns 1, 2, 3 and 4) and the results of the measurements (see below).

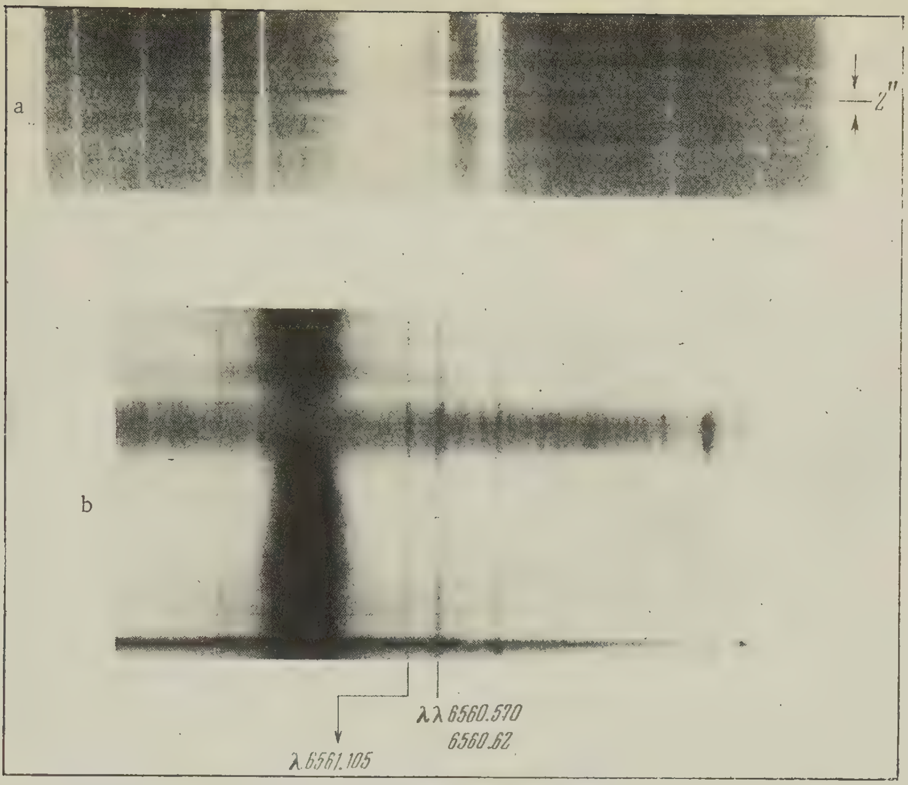


Fig. 1. a) Emission "moustaches" near the H_α-line. b) The strengthening of the line λ 6561.105 A in the region of the sunspot of January 4, 1957.

All the spectrograms (except Nos. 617 and 873) were obtained with the large spectrograph of the Solar Tower Telescope (STT) at a resolving power of 300,000 (the instrumental width is 0.02 A) and a dispersion of 4 mm = = 1 A, with the help of a 150 · 150 grating which concentrates light in the fourth order [4]. (The use of photoelectric methods is impossible here because of the very small dimensions of the emission granules investigated — about 2-5".) The spectrograms 617 and 873 were obtained at a resolving power and dispersion one-half as large. Examples of a microphotogram of part of the H_α blue wing containing the line λ 6561.105 A are given in Figure 2 for the region of the "moustaches" and for the undisturbed atmosphere in its vicinity. The profile of the line λ 6561.105 A (actually, the region containing this line and the neighboring combination of λ 6560.570 and 6560.62) was determined in units of the residual intensity of this wing; the equivalent width and the residual intensity at the center of this line were measured for the disturbed (W and $r_\nu(0)$) and the neighboring undisturbed region (W_0 and $r_{\nu_0}(0)$). The results of the measurements are given in Table 1 (Columns 5, 6, 8 and 9). The magnitude of the relative change in the equivalent width $\Delta W/W_0$ and of the change in the residual intensities are also given in the Table (Columns 7 and 10). The analysis of these two sets of data shows that in 21 of the 28 cases the equivalent width of the line λ 6561.105 A increases for the region of the emission granule; only in 7 cases it decreases. The residual intensity of this line decreases in 15 cases, increases in 4 cases, and remains unchanged in 9 cases. The average increase of the equivalent width (for all 28 spectra) is

$$\frac{\overline{\Delta W}}{W_0} = +0.319, \tag{1}$$

and the decrease of the residual central intensity is

$$\overline{\Delta r_\nu} = -1.7\,\%. \tag{2}$$

TABLE 1

No.	Date and phenomenon observed	Spectrogram No.	Spectrum No.	$\lambda = 6561{\cdot}105$					
				W*	W_0	$\Delta W/W_0$	r_ν (0)	$r_{\nu 0}(0)$	Δr_ν (%)
1	Aug. 25, 1955, "moustaches"	617	1	53.1	40.3	+0.32	0.72	0.74	—2
2	"		2	36.3	43.2	—0.16	0.74	0.74	0
3	"		4	49.9	42.9	+0.16	0.72	0.72	0
4	Oct. 17, 1955		1	36.7	18.4	+0.99	0.83	0.90	—7
5	"moustaches"	873	2	27.9	18.1	+0.53	0.84	0.91	—7
6	"		4	46.2	18.9	+1.44	0.83	0.82	+1
7	May 15, 1956, "moustaches"	1052a	1	5.3	4.9	+0.08	0.95	0.95	0
8	June 22, 1956, "moustaches"	1073	1	25.9	26.9	—0.04	0.84	0.835	+0.5
9	Flare, "moustaches"		3	29.4	35.7	—0.18	0.82	0.81	+1
10			4	28.6	23.0	+0.24	0.83	0.84	—1
11	July 6, 1956	1108	1	13.5	8.9	+0.52	0.94	0.96	—2
12	"moustaches",		3	9.4	9.1	+0.03	0,94	0.94	0
13	"		4	15.3	7.3	+1,09	0.925	0.95	—0.25
14	"		5	9.3	12.9	—0.28	0.94	0.94	0
15	Aug. 12, 1956 flare	1122	4	28.9	15.3	+0.64	0.85	0.92	—7
16	"		6	29.9	22.4	+0.34	0.87	0.90	—3
17	"		10	34.4	22.1	+0.56	0.87	0.90	—3
18	July 30, 1956	1175	1	10.9	6.0	+0.82	0.92	0.93	—1
19	Aug. 12, 1956	1248	1	15.8	18.1	—0.13	0.89	0.89	0
20	"moustaches"	1287	1	24.1	17.3	+0.39	0.88	0.89	—1
21	"	1288	1	25.1	21.7	+0.16	0.85	0.84	+1
22	"		2	24.8	25.9	—0.04	0.84	0.84	0
23	"		3	26.2	23.8	+0.10	0.84	0.84	0
24	Aug. 21, 1956 cont. em.	1290	1	22.4	11.0	+1.14	0,88	0.92	—4
25	Sept. 7, 1956, "moustaches"	1438	1	59.7	58.2	+0.03	0.68	0.69	—1
26	"		3	60.0	51.9	+0.16	0.66	0.74	—8
27	"		4	43.5	40.7	+0.07	0.76	0.77	—7
28	"		5	47.1	47.3	—0.04	0.78	0.78	0
						Av. $\frac{\Delta W}{W_0} = +0.319$			Av. $\Delta r_\nu = -1.7\%$

---

* The unit of W is 0.8 mA (except Nos. 617 and 873, where it is 0.4 mA).

TABLE 2

Spectrogram No.	No. of spectra	n	$\overline{W_0}$	ε'_{W_0}	ε_{W_0}	$\overline{r_{\nu_0}(0)}$	$\varepsilon\,[r_{\nu_0}(0)]$
1073	3	9	22.1	2.36	0.107	0.846	1.41%
1108	3	9	9.6	1.22	0.127	0.934	0.76
1122	3	15	17.4	3.78	0.212	0.887	1.18
1290	2	6	17.9	2.09	0.117	0.877	0.98
		Total 39			Av ±0.141		Av ±1.08%

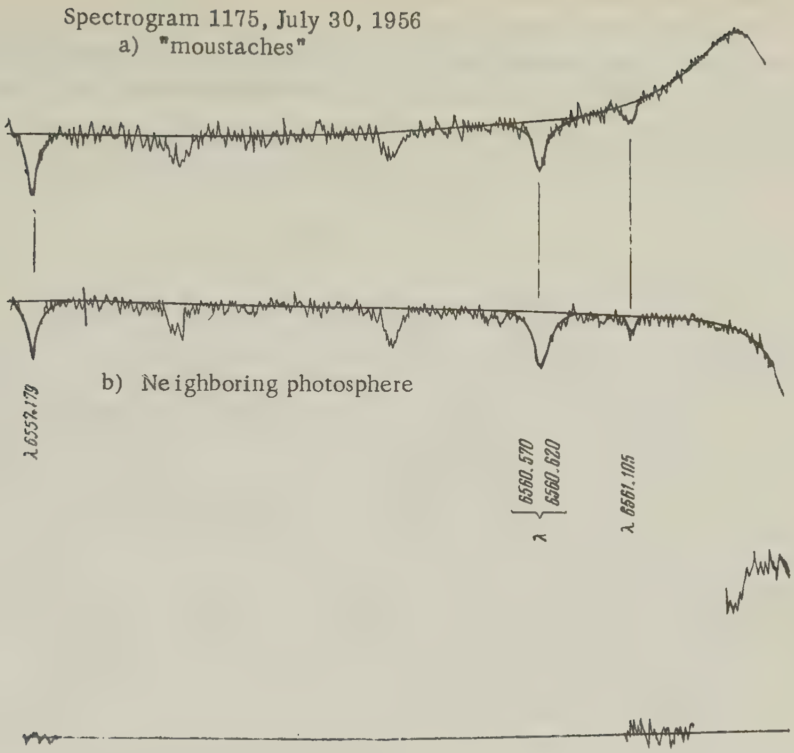


Fig. 2. Microphotometer tracings of the H_α blue wing for the region of the "moustaches" (above) and for the undisturbed neighboring atmosphere (below).

However, in order to determine whether this result is real or not, first of all it is necessary to determine the extent to which it is affected by random errors. For this purpose, 39 tracings from the same plates were made in the same spectral region for the undisturbed solar atmosphere, away from the regions of emission granules and of sunspots. Three cases showing a large increase of $\Delta W/W_0$ (Nos. 1122, 1290, 1108) and one not showing any increase (1073) were chosen; these cases also corresponded to various amounts of water vapor in the earth's atmosphere, as could be seen from the intensity of the investigated line λ 6561.105 A and the other atmospheric lines. In each case the average profile of λ 6561.105 A was determined and for each point on this profile, including the central residual intensity, the mean square error of an individual measurement $\epsilon(r_\nu)$ was found; W_0, the average $\overline{W}_0$, and the mean square error of one measurement ϵ'_{W_0} were determined for each profile. The results are given in Table 2, where n, the number of profiles measured, was determined by the height of the spectrum available for measurement; the number of spectra is given in Column 2; $\overline{r_{\nu_0}(0)}$ is the average central residual intensity, and $\epsilon_{W_0} = \epsilon'_{W_0}/\overline{W}_0$ (in view of the day-to-day variations of the intensity of the water-vapor lines, the accuracy of their equivalent-width measurements should be characterized by the magnitude of the relative error). The mean errors ϵ, associated with differences in the exposure and processing of the spectrograms (water-vapor concentration, transparency of the atmosphere, photographic density, effect of emulsion grains, errors of measurement and tracing), in general determine the accuracy of individual measurements of W_0 and $r_\nu(0)$ for lines with a Rowland intensity from -2 to 0 in our spectrograms.

Figure 3 (on the left) gives the frequency distribution of the quantity $\Delta W/W_0$, which is of interest to us, according to the data of Table 1 (full line, points) and the frequency distribution of $\Delta W_0/W_0$, the errors of the 39 individual measurements of the equivalent width which were used to construct Table 2 (dashed line, circles). Figure 4 gives the same distribution for the quantities $\Delta r_\nu(0)$ (full line, points) and $\Delta r_{\nu_0}(0)$ (dashed line, circles). From Figure 3 it can be seen that while the distribution of negative values of $\Delta W/W_0$ practically coincides with the normal distribution, the distribution of positive values of this quantity deviates from the normal distribution of errors in the determination of this quantity to such an extent that it provides evidence for the assumption that the increase of the equivalent width of λ 6561.105 in the active regions is not due to chance. Even if it were valid for us to reject all positive values of this quantity greater than the arithmetic mean (i. e., about one-half of the total), then in this case also there would be some deviation from the normal distribution and an excess of positive values of $\Delta W/W_0$. We have a similar situation also for the distribution of Δr_ν values, but in this case the deviation is less marked because of the high uncertainty of central-intensity measurements in general.

Therefore, the increase of the equivalent width of λ 6561.105 A for emission granules cannot be the result of random errors. Let us now investigate the influence of the systematic errors. If the line λ 6561.105 A were purely an atmospheric line due to water vapor, then its behavior in regions of continuous emission would not be different in any way from the behavior of other atmospheric lines, i. e., its equivalent width would remain the same (within the limits of measurement errors), irrespective of the position on the solar disc. The fact that the equivalent width of λ 6561.105 A increases in the emission regions already excludes this possibility. However, unlikely as it is, let us find out if it is possible for some changes in the strength of telluric lines to occur in going

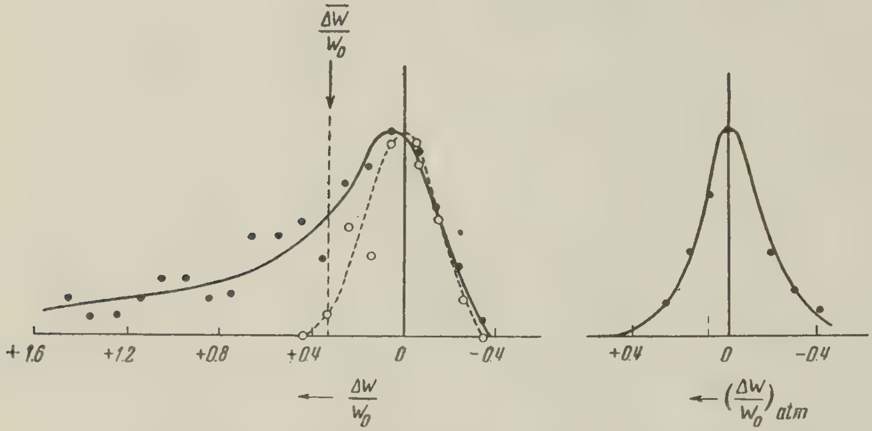


Fig. 3. On the left is the frequency distribution of $\Delta W/W_0$ for λ 6561.105 (dots, full line) and the distribution of the errors of the equivalent-width measurements (circles, dashed line). The dashed vertical line is the mean value. On the right — the distribution of $\Delta W/W_0$ for the atmospheric lines λ 6558.162 and λ 6557.179 A.

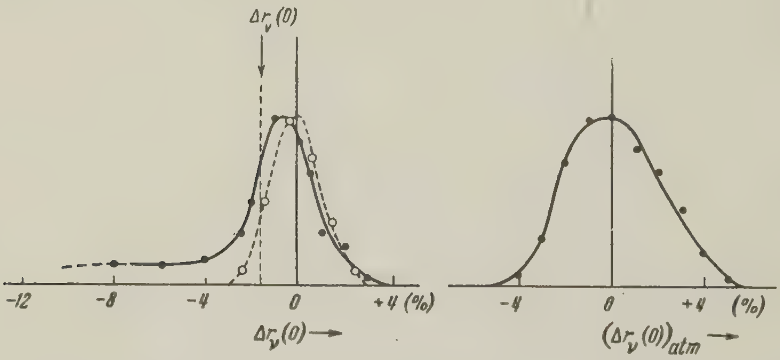


Fig. 4. On the left is the frequency distribution of $\Delta r_\nu(0)$ (dots, full line) and the distribution of the errors of the measurement of the central residual intensity $\Delta r_{\nu_0}(0)$ (circles, dashed line). On the right is the distribution of $\Delta r_{\nu_0}(0)$ for the atmospheric lines λ 6558.162 and λ 6557.179 A.

from an undisturbed region on the sun to the emission regions (for example, because of effects due to a high photographic density, the tremor of the emission granule on the spectrograph slit, and other possible reasons). For this purpose the profiles of other atmospheric lines, λ 6558.162 A and λ 6557.179 A, were obtained for the emission granules and outside them (in most cases both lines were obtained from the same spectrograms as the line λ 6561.105 under investigation, see Fig. 2), in the case of the spectrograms listed in Table 1. The distribution curves of $\Delta W/W_0$ and $\Delta r_\nu(0)$ for these lines are given on the right of Figures 3 and 4, respectively. We observe that both these distributions are very close to the normal distribution; in this case the average values are

$$\overline{\frac{\Delta W}{W_0}} = -0.005; \qquad \overline{\Delta r_\nu(0)} = +0.7\%;$$

i. e., these lines show no appreciable change of the equivalent width or of the central intensity. On the contrary, they exhibit the opposite effect – the strengths of these lines are somewhat decreased as compared with the lines of the emission granules, as was to be expected because 1) the image of the lines is flooded as the result of the high photographic density of the emission-granule spectrum, and 2) there is a small tremor of the line image in the spectrograph when air currents are present inside the building. It must be noted that the effect of the tremor of the emission granule or the sunspot on the spectrograph slit must always decrease the difference (if it exists) between the line profiles for the active region and the photosphere.

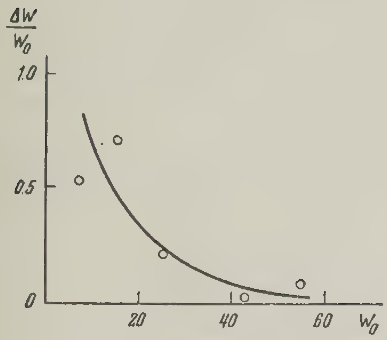


Fig. 5. The increase of the equivalent width of λ 6561.105 A as a function of the intensity of this line.

Therefore, the investigation of the equivalent width and the central intensity of λ 6561.105 A indicates the presence of some weak solar line which blends with the water-vapor line and which is strengthened in the emission granules as compared with the undisturbed solar atmosphere. It is obvious that the smaller the concentration of water vapor in the earth's atmosphere and the weaker the line λ 6561.105 A, the more chances are there to observe this strengthening. This is confirmed by the dependence of $\Delta W/W_0$ on W_0 shown in Figure 5, which, in addition, indicates that the increase of these quantities is real. For large values of W_0 and $r_\nu(0)$ it is completely impossible to observe the presence of this solar line. This is the main reason* why Kinnman [6] was not successful in his attempt to observe the D_α-line, since he worked in Oxford under conditions of a high atmospheric humidity (the depth of the line λ 6561.105 A in his observations was as large as 20%).

However, the decisive evidence for the presence of a solar line with a wavelength close to λ 6561.105 A comes from spectrograms of the H_α region near to a sunspot which were recently obtained with the Solar Tower Telescope under very exceptional conditions on January 4, 1957 and January 23, 1957: never before have we seen a spectrogram in which the atmospheric lines of water vapor were so weak that their presence could only be guessed at. One of these spectrograms is reproduced in Figure 1 (lower part): we see that the line λ 6561.105 A is barely visible outside the sunspot spectrum, while in the sunspot spectrum it is clearly visible as a fairly strong line. However, inasmuch as this judgement can be influenced by the contrast, we give in Figure 6 the microphotometer tracing of the blue H_α-wing for the spot (lower tracing) and for a neighboring region of the photosphere (middle tracing) in the same spectrogram; at the top, we give for comparison a tracing of the same region of the spectrum for a typical day with a moderately high humidity. We see that on these exceptional days, January 4 and 23, 1957, the line λ 6561.105 A was practically absent from the spectrum of the undisturbed solar disc; further, the water-vapor line λ 6557.179 A, which is at least three times as strong as λ 6561.105, was also barely visible on this tracing. At the same time, a fairly strong line, whose intensity is comparable with that of λ 6569.588 Ti^+ appears in the sunspot spectrum (lower tracing) in place of the line λ 6561.105 A. The wavelength of this line, obtained from measurements relative to the neighboring weak solar lines, is

$$\lambda = 6561.09,$$

---

*As well as the relatively low resolving power of Kinnman's method (~ 100,000).

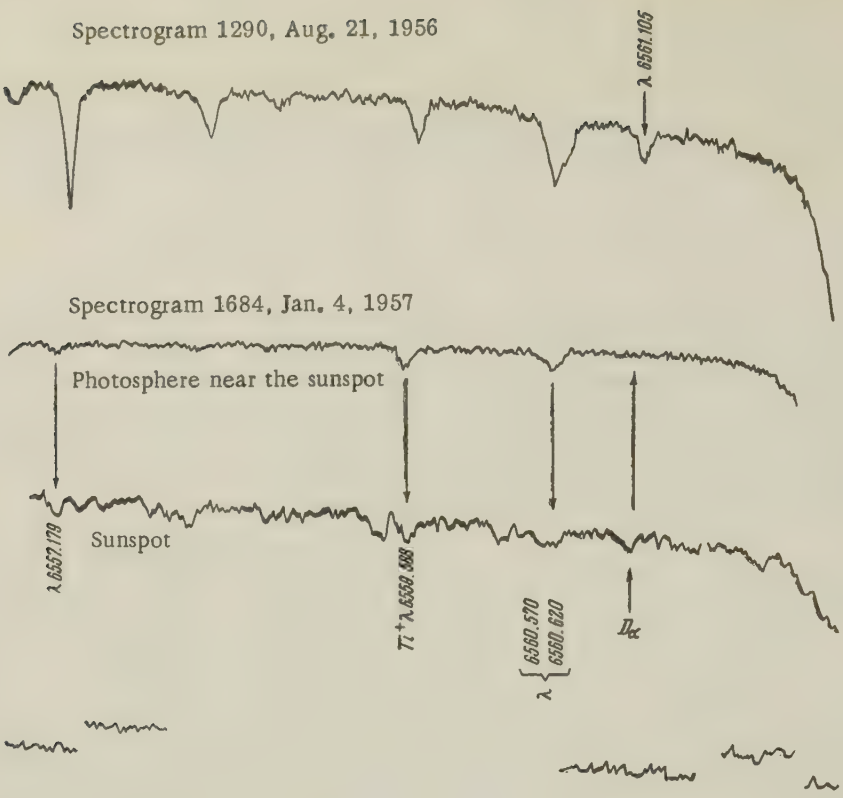


Fig. 6. Microphotometer tracings of the violet wing of the H_α-line for days with an exceptionally low water vapor concentration in the atmosphere — January 4 and 23, 1957; lower tracing — for the region of the sunspot, middle tracing — for the neighborhood of the sunspot in the same spectrogram. At the top (for comparison) — the same region of the spectrum for days with a moderately high content of water vapor in the atmosphere.

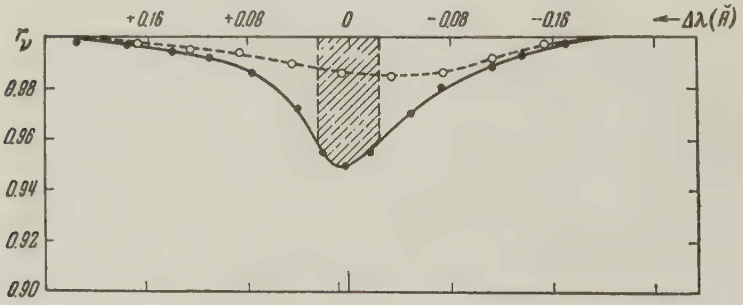


Fig. 7. The profile of the line λ 6561.105 A for the sunspot (dots, full line) and for the neighboring photosphere (circles, dashed line).

practically coinciding with the wavelength of the red component of the Dα-line fine structure (see [1]). The line is also seen to be asymmetric: its violet wing is appreciably flatter than the red wing, which is to be expected for the Dα-line (see [1]). In spite of the weakness of the line λ 6561.105 A in the spectra of the regions in the vicinity of the sunspot, we were able in some spectrograms to determine its profile and to compare it with the profile of the solar line λ 6561.09 A in the sunspot spectrum (Fig. 7, where the dots denote the profile of λ 6561.09 A and the circles the profile of λ 6561.105 A for the neighboring photosphere; the width of the strip along the center of the line corresponds to twice the width of the instrumental profile). The equivalent widths of these two lines:

$$W(\lambda\ 6561.105) = 3.4\ \text{mA},$$

$$W(\lambda\ 6561.09) = 6.4\ \text{mA},$$

show that the increase of the equivalent width of the solar line in going from the photosphere to the region of the sunspot is at least 3 mA, which corresponds to an increase of the relative concentration of deuterium (if this is a deuterium line) to a value $D/H \sim 10^{-4}$. Similar results are obtained from the spectrograms of January 23, 1957.

Therefore, it is highly probable that the solar line appearing very close to the line λ 6561.105 A is the D_α-line and that it is strengthened in the emission granules and sunspots. Since at the present time we do not know of any line, except D_α, which has a wavelength so close to that of this line, it seems natural to conclude that there is a strengthening of the deuterium line in the active formations on the sun. If this is so, we are forced to conclude that unusual nuclear reactions which lead to an increase of the deuterium concentration, occur in the active formations on the solar surface. This can hardly be the result of thermonuclear reactions with hydrogen being carried out of the deeper regions of the sun, since the deuterium must already be almost completely burned-up at temperatures above $5 \cdot 10^5$°K. It is possible that there is an increase of neutron production in the active regions on the solar surface, analogous to the neutron production in plasmas as was observed in the well-known experiments of L. A. Artsimovich and others [7]. The production of cosmic rays in the active regions on the sun indicates that processes liberating large amounts of energy take place on the solar surface. However, these problems, which require the attention of astrophysicists, need very careful investigations in the future.

Crimean Astrophysical Observatory
of the USSR Academy of Sciences

Received April 5, 1957

LITERATURE CITED

[1] A. B. Severnyi, Izv. Krymskoi Astrofiz. Obs. 16, 12 (1956).

[2] A. B. Severnyi, Astron. J. (USSR) 33, 74 (1956); see also [3].

[3] A. B. Severnyi, Izv. Krymskoi Astrofiz. Obs. 17, 102 (1957).

[4] A. B. Severnyi, Izv. Krymskoi Astrofiz. Obs. 15, 31 (1955).

[5] T. Kinnman, Monthly Notices, Roy. Astron. Soc. 166, 77 (1956).

[6] L. A. Artsimovich, J. Atomic Energy (USSR) 1956, No. 3.

THE DIFFUSION OF RADIATION IN A MEDIUM OF FINITE OPTICAL THICKNESS

V. V. Sobolev

The problem of the diffusion of radiation in a medium of finite optical thickness τ_0 and with a spherical indicatrix of scattering is considered. In Section 1 the equations for Ambartsumian's functions $\varphi(\eta, \tau_0)$ and $\psi(\eta, \tau_0)$ are derived by a new method. In Section 2 the asymptotic formulas for these functions for large optical thickness ($\tau_0 >> 1$) are given. The expressions for the intensity of radiation emerging from a medium are derived in Section 3 for various radiation sources.

In one of his previous papers [1], the author proposed a new method for the solution of various problems in the theory of radiation diffusion which is based on the use of the probability for escape of a quantum from the medium. This method was first applied to the case of a semi-infinite medium. Then, in reference [2], it was also applied to the case of a medium with a finite optical thickness. In the present paper some new results are obtained by this method for a medium of finite optical thickness. Particular attention is paid to the case when the optical thickness is very large.

The results obtained here can be used in the study of the diffusion of radiation in gaseous nebulae, stellar envelopes, planetary atmospheres, and other objects.

1. Fundamental Equations

Let the medium consist of plane-parallel layers and have an optical thickness τ_0. We will assume that the scattering in the medium is spherically symmetric. The probability that the quantum survives an elementary act of scattering (or the albedo of the particle) we will denote by λ. The intensity of the radiation sources we will consider to be a function of only the optical depth τ. During the diffusion of the radiation in the medium, some of the quanta escape from it, while the others suffer true absorption in the medium (if $\lambda < 1$). The problem consists in determining the intensity of the radiation escaping from the medium.

To solve the problem presented, let us introduce the probability for the escape of the quanta from the medium. Let $p(\tau, \eta, \tau_0)$ be the probability that a quantum produced at an optical thickness τ escapes from the medium (in general after diffusing through it) through the boundary $\tau = 0$, into the element of solid angle $d\omega$ at an angle arc cos η to the normal. It is obvious that if the function $p(\tau, \eta, \tau_0)$ is known, the intensity of the radiation escaping from the medium can be easily found for any energy sources acting in the medium.

In order to obtain the equations for the function $p(\tau, \eta, \tau_0)$, we proceed as follows. Let us find the probability for the escape of a quantum from the depth $\tau + \Delta\tau$, i. e., the quantity $p(\tau + \Delta\tau, \eta, \tau_0)$, assuming that $\Delta\tau$ is a small quantity whose square can be neglected. The escape of the quantum from the depth $\tau + \Delta\tau$ in a medium of optical thickness τ_0 can be considered as the escape of the quantum from the depth τ in a medium of optical thickness $\tau_0 - \Delta\tau$ and the subsequent traversal by the quantum of the layer of thickness $\Delta\tau$ (with or without absorption in this layer). We therefore have

$$p(\tau+\Delta\tau, \eta, \tau_0) = p(\tau, \eta, \tau_0-\Delta\tau)\left(1-\frac{\Delta\tau}{\eta}\right)+ + 2\pi\int_0^1 p(\tau, \eta', \tau_0)\frac{\Delta\tau}{\eta'}d\eta' p(0, \eta, \tau_0). \tag{1}$$

It follows from this that

$$\frac{\partial p}{\partial\tau} = -\frac{\partial p}{\partial\tau_0} - \frac{1}{\eta}p(\tau, \eta, \tau_0) + 2\pi\int_0^1 p(\tau, \eta', \tau_0)\frac{d\eta'}{\eta'}p(0, \eta, \tau_0). \tag{2}$$

To determine $\partial p/\partial\tau_0$ let us find the quantity $p(\tau, \eta, \tau_0-\Delta\tau)$. It is obvious that

$$p(\tau, \eta, \tau_0-\Delta\tau) = p(\tau, \eta, \tau_0) - 2\pi\int_0^1 p(\tau_0-\tau, \eta', \tau_0)\frac{\Delta\tau}{\eta'}d\eta' p(\tau_0, \eta, \tau_0). \tag{3}$$

This yields

$$\frac{\partial p}{\partial\tau_0} = 2\pi\int_0^1 p(\tau_0-\tau, \eta', \tau_0)\frac{d\eta'}{\eta'}p(\tau_0, \eta, \tau_0). \tag{4}$$

The functions $p(0, \eta, \tau_0)$ and $p(\tau_0, \eta, \tau_0)$ appearing in Equations (2) and (4) can be written in the form:

$$p(0, \eta, \tau_0) = \frac{\lambda}{4\pi}\left[1 + 2\pi\int_0^1 d\zeta\int_0^{\tau_0} e^{-\tau/\zeta}p(\tau, \eta, \tau_0)\frac{d\tau}{\zeta}\right]; \tag{5}$$

$$p(\tau_0, \eta, \tau_0) = \frac{\lambda}{4\pi}\left[e^{-\tau_0/\eta} + 2\pi\int_0^1 d\zeta\int_0^{\tau_0} e^{-(\tau_0-\tau)/\zeta}p(\tau, \eta, \tau_0)\frac{d\tau}{\zeta}\right]. \tag{6}$$

Let us introduce the quantities:

$$\rho(\eta, \zeta, \tau_0) = \frac{\pi}{\eta\zeta}\int_0^{\tau_0} e^{-\tau/\zeta}p(\tau, \eta, \tau_0)\,d\tau; \tag{7}$$

$$\sigma(\eta, \zeta, \tau_0) = \frac{\pi}{\eta\zeta}\int_0^{\tau_0} e^{-(\tau_0-\tau)/\zeta}p(\tau, \eta, \tau_0)\,d\tau; \tag{8}$$

$$p(0, \eta, \tau_0) = \frac{\lambda}{4\pi}\varphi(\eta, \tau_0), \quad p(\tau_0, \eta, \tau_0) = \frac{\lambda}{4\pi}\psi(\eta, \tau_0). \tag{9}$$

Then, instead of (5) and (6) we have

$$\varphi(\eta, \tau_0) = 1 + 2\eta\int_0^1 \rho(\eta, \zeta, \tau_0)\,d\zeta; \tag{10}$$

$$\psi(\eta, \tau_0) = e^{-\tau_0/\eta} + 2\eta\int_0^1 \sigma(\eta, \zeta, \tau_0)\,d\zeta. \tag{11}$$

The physical meaning of the functions $\varphi(\eta, \tau_0)$ and $\psi(\eta, \tau_0)$ can be seen from Equation (9). The quantities $\rho(\eta, \zeta, \tau_0)$ and $\sigma(\eta, \zeta, \tau_0)$ are the coefficients of brightness for a medium illuminated by parallel rays (see Section 3).

Using Equations (2), (4), (10) and (11) we can express the quantities $\rho(\eta,\zeta,\tau_0)$ and $\sigma(\eta,\zeta,\tau_0)$ in terms of the functions $\varphi(\eta,\tau_0)$, $\psi(\eta,\tau_0)$ and obtain the equations for determining these functions.

Multiplying (2) by $e^{-\tau/\zeta}$ and integrating with respect to τ within the limits from 0 to τ_0, we have

$$\begin{aligned} -p(0,\eta,\tau_0) + \frac{1}{\zeta}\int_0^{\tau_0} e^{-\tau/\zeta} p(\tau,\eta,\tau_0)\,d\tau = \\ = -\frac{\partial}{\partial\tau_0}\int_0^{\tau_0} e^{-\tau/\zeta} p(\tau,\eta,\tau_0)\,d\tau - \frac{1}{\eta}\int_0^{\tau_0} e^{-\tau/\zeta} p(\tau,\eta,\tau_0)\,d\tau + \\ + 2\pi\int_0^1 \frac{d\eta'}{\eta'}\int_0^1 e^{-\tau/\zeta} p(\tau,\eta',\tau_0)\,d\tau\, p(0,\eta,\tau_0) \end{aligned} \tag{12}$$

or, taking (7), (9) and (10) into account,

$$(\eta+\zeta)\,\rho(\eta,\zeta,\tau_0) + \eta\zeta\frac{\partial\rho}{\partial\tau_0} = \frac{\lambda}{4}\,\varphi(\eta,\tau_0)\,\varphi(\zeta,\tau_0). \tag{13}$$

Multiplying (2) by $e^{-(\tau_0-\tau)/\zeta}$, integrating with respect to τ from 0 to τ_0, and using (8), (9) and (11), we obtain

$$\zeta\sigma(\eta,\zeta,\tau_0) + \eta\zeta\frac{\partial\sigma}{\partial\tau_0} = \frac{\lambda}{4}\,\varphi(\eta,\tau_0)\,\psi(\zeta,\tau_0). \tag{14}$$

Similarly, from Equation (4) we find

$$\eta\zeta\frac{\partial\rho}{\partial\tau_0} = \frac{\lambda}{4}\,\psi(\eta,\tau_0)\,\psi(\zeta,\tau_0), \tag{15}$$

$$\eta\sigma(\eta,\zeta,\tau_0) + \eta\zeta\frac{\partial\sigma}{\partial\tau_0} = \frac{\lambda}{4}\,\psi(\eta,\tau_0)\,\varphi(\zeta,\tau_0). \tag{16}$$

From the relations presented above, we are able to obtain three systems of equations for determining the functions $\varphi(\eta,\tau_0)$ and $\psi(\eta,\tau_0)$.

1) Subtracting (15) from (13) and (16) from (14), we find

$$\rho(\eta,\zeta,\tau_0) = \frac{\lambda}{4}\,\frac{\varphi(\eta,\tau_0)\,\varphi(\zeta,\tau_0) - \psi(\eta,\tau_0)\,\psi(\zeta,\tau_0)}{\eta+\zeta}; \tag{17}$$

$$\sigma(\eta,\zeta,\tau_0) = \frac{\lambda}{4}\,\frac{\varphi(\eta,\tau_0)\,\psi(\zeta,\tau_0) - \psi(\eta,\tau_0)\,\varphi(\zeta,\tau_0)}{\zeta-\eta}. \tag{18}$$

The substitution of (17) into (10) and of (18) into (11) gives

$$\varphi(\eta,\tau_0) = 1 + \frac{\lambda}{2}\,\eta\int_0^1 \frac{\varphi(\eta,\tau_0)\,\varphi(\zeta,\tau_0) - \psi(\eta,\tau_0)\,\psi(\zeta,\tau_0)}{\eta+\zeta}\,d\zeta; \tag{19}$$

$$\psi(\eta,\tau_0) = e^{-\tau_0/\eta} + \frac{\lambda}{2}\,\eta\int_0^1 \frac{\varphi(\eta,\tau_0)\,\psi(\zeta,\tau_0) - \psi(\eta,\tau_0)\,\varphi(\zeta,\tau_0)}{\zeta-\eta}\,d\zeta. \tag{20}$$

2) From Equations (15) and (16) it follows that

$$\rho(\eta,\zeta,\tau_0) = \frac{\lambda}{4}\int_0^{\tau_0} \psi(\eta,\tau)\,\psi(\zeta,\tau)\,\frac{d\tau}{\eta\zeta}, \tag{21}$$

$$\sigma(\eta, \zeta, \tau_0) = \frac{\lambda}{4}\int_0^{\tau_0} e^{-(\tau_0-\tau)/\zeta}\, \psi(\eta, \tau)\, \varphi(\zeta, \tau)\, \frac{d\tau}{\eta\zeta}. \tag{22}$$

Substituting (21) into (10) and (22) into (11), we find

$$\varphi(\eta, \tau_0) = 1 + \frac{\lambda}{2}\int_0^1 \frac{d\zeta}{\zeta}\int_0^{\tau_0} \psi(\eta, \tau)\, \psi(\zeta, \tau)\, d\tau, \tag{23}$$

$$\psi(\eta, \tau_0) = e^{-\tau_0/\eta} + \frac{\lambda}{2}\int_0^1 \frac{d\zeta}{\zeta}\int_0^{\tau_0} e^{-(\tau_0-\tau)/\zeta}\, \psi(\eta, \tau)\, \varphi(\zeta, \tau)\, d\tau. \tag{24}$$

3) From (13) and (14) we have

$$\rho(\eta, \zeta, \tau_0) = \frac{\lambda}{4}\int_0^{\tau_0} e^{-(\tau_0-\tau)\left(\frac{1}{\eta}+\frac{1}{\zeta}\right)}\, \varphi(\eta, \tau)\, \varphi(\zeta, \tau)\, \frac{d\tau}{\eta\zeta}; \tag{25}$$

$$\sigma(\eta, \zeta, \tau_0) = \frac{\lambda}{4}\int_0^{\tau_0} e^{-(\tau_0-\tau)/\eta}\, \varphi(\eta, \tau)\, \psi(\zeta, \tau)\, \frac{d\tau}{\eta\zeta}. \tag{26}$$

The substitution of (25) into (10) and of (26) into (11) yields the following equations:

$$\varphi(\eta, \tau_0) = 1 + \frac{\lambda}{2}\int_0^1 \frac{d\zeta}{\zeta}\int_0^{\tau_0} e^{-(\tau_0-\tau)\left(\frac{1}{\eta}+\frac{1}{\zeta}\right)}\, \varphi(\eta, \tau)\, \varphi(\zeta, \tau)\, d\tau; \tag{27}$$

$$\psi(\eta, \tau_0) = e^{-\tau_0/\eta} + \frac{\lambda}{2}\int_0^1 \frac{d\zeta}{\zeta}\int_0^{\tau_0} e^{-(\tau_0-\tau)/\eta}\, \varphi(\eta, \tau)\, \psi(\zeta, \tau)\, d\tau. \tag{28}$$

Equations (17), (18) and (19), (20) were first obtained by V. A. Ambartsumian [3]. Equations (21), (22) and (23), (24) were also obtained by him [2]. The fundamental integral equation of the theory of radiation diffusion was used to derive them. Equations (25), (26) and (27), (28), as far as we know, have not been derived before.*

Thus, using the method proposed by us, we were able to obtain very simply the known, as well as simultaneously the new, equations for the functions $\varphi(\eta, \tau_0)$ and $\psi(\eta, \tau_0)$.

After the functions $\varphi(\eta, \tau_0)$ and $\psi(\eta, \tau_0)$ have been found, the probability for the escape of the quantum from the medium $p(\tau, \eta, \tau_0)$ can be determined from the equation:

$$\begin{aligned}\frac{\partial p}{\partial \tau} = &-\frac{1}{\eta}\, p(\tau, \eta, \tau_0) + \frac{\lambda}{2}\, \varphi(\eta, \tau_0)\int_0^1 p(\tau, \eta', \tau_0)\, \frac{d\eta'}{\eta'} - \\ &- \frac{\lambda}{2}\, \psi(\eta, \tau_0)\int_0^1 p(\tau_0-\tau, \eta', \tau_0)\, \frac{d\eta'}{\eta'},\end{aligned} \tag{29}$$

which is obtained from Equations (2), (4) and (9). However, we will not proceed to solve Equation (29) here, since in many cases the intensity of the radiation emerging from the medium can be found from Equation (29) without an explicit knowledge of the function $p(\tau, \eta, \tau_0)$.

2. Asymptotic Expressions for the Functions $\varphi(\eta, \tau_0)$ and $\psi(\eta, \tau_0)$

The above equations for the functions $\varphi(\eta, \tau_0)$ and $\psi(\eta, \tau_0)$ can be solved numerically. Considerable progress in this direction has been made by Chandrasekhar, Elbert and Franklin [5] who, by solving Equations (19) and (20) numerically, have compiled detailed tables of the functions $\varphi(\eta, \tau_0)$ and $\psi(\eta, \tau_0)$ for $\lambda = 1, 0.95, 0.9, 0.8, 0.5$ and for values of τ_0 between 0 and 1.

---

*Equations (25) and (27) can be deduced from the more general equations derived by the author [4] for the case of a medium in which λ depends on τ.

Recently, Busbridge has obtained linear equations for Equations (19) and (20) (these equations were earlier obtained by the author [7] using a different method) and has determined some of the properties of $\varphi(\eta, \tau_0)$ and $\psi(\eta, \tau_0)$. In particular, she proved that a solution of these equations exists for large values of τ_0.

In the present section we derive, with the help of the above equations, an asymptotic expression for the functions $\varphi(\eta, \tau_0)$ and $\psi(\eta, \tau_0)$ when the optical thickness is large. Initially, we investigate the case when, in addition to scattering, there is true absorption of the radiation in the medium ($\lambda < 1$), then the case when there is only scattering in the medium ($\lambda = 1$).

First of all we note that if the optical thickness of the medium is infinitely large, $\psi(\eta, \tau_0) = 0$ while $\varphi(\eta, \tau_0)$ becomes the function $\varphi(\eta)$ given by the equation

$$\varphi(\eta) = 1 + \frac{\lambda}{2}\eta\varphi(\eta)\int_0^1 \frac{\varphi(\zeta)}{\eta+\zeta}d\zeta. \tag{30}$$

Equation (30) has been studied in detail. Accurate and approximate expressions, as well as tables, have been published for the function $\varphi(\eta)$. In the following we will consider $\varphi(\eta)$ as a known function.

The case when $\lambda < 1$. We will investigate the system of Equations (27) and (28) assuming that the optical thickness of the medium is very large ($\tau_0 \gg 1$).

We will seek a solution in the form:

$$\varphi(\eta, \tau_0) = \varphi(\eta) - \chi(\eta)e^{-2k\tau_0}, \tag{31}$$

where k is a constant. Substituting (31) into (27), neglecting terms smaller than $e^{-2k\tau_0}$, and using (30), we find

$$\chi(\eta) = \frac{\lambda}{2}\eta\int_0^1 [\varphi(\eta)\chi(\zeta) + \varphi(\zeta)\chi(\eta)]\frac{d\zeta}{\eta+\zeta-2k\eta\zeta}. \tag{32}$$

The solution of Equation (32) is

$$\chi(\eta) = C\frac{\eta}{1-k\eta}\varphi(\eta), \tag{33}$$

where λ and k are linked by the expression

$$\frac{\lambda}{2}\int_0^1 \frac{\varphi(\eta)}{1-k\eta}d\eta = 1, \tag{34}$$

and C is a constant. Using (30), we can obtain the following expression instead of (34) (Reference [2], p. 114):

$$\frac{\lambda}{2k}\log\frac{1+k}{1-k} = 1. \tag{35}$$

Thus, for large values of τ_0 the function $\varphi(\eta, \tau_0)$ is given by

$$\varphi(\eta_0\tau_0) = \varphi(\eta) - C\frac{\eta}{1-k\eta}\varphi(\eta)e^{-2k\tau_0}. \tag{36}$$

Similarly, from Equation (28) we obtain the following asymptotic expression for the function $\psi(\eta, \tau_0)$:

$$\psi(\eta, \tau_0) = C_1 \frac{\eta}{1 - k\eta} \varphi(\eta) e^{-k\tau_0}, \tag{37}$$

where C_1 is a new constant.

To determine the constants C and C_1 we proceed as follows. Let us find, with the help of Equation (29), the quantities

$$\begin{aligned} A(m, \eta, \tau_0) &= \frac{4\pi}{\lambda} \int_0^{\tau_0} p(\tau, \eta, \tau_0)\, e^{-m(\tau_0 - \tau)}\, d\tau, \\ B(m, \eta, \tau_0) &= \frac{4\pi}{\lambda} \int_0^{\tau_0} p(\tau, \eta, \tau_0)\, e^{-m\tau}\, d\tau \end{aligned} \tag{38}$$

(in what follows we will require them for another reason).

Multiplying (29) by $e^{m\tau}$, integrating with respect to τ from 0 to τ_0, and then using (9), we get

$$\begin{aligned} \left(\frac{1}{\eta} - m\right) A(m, \eta, \tau_0) &= \varphi(\eta, \tau_0) \left[e^{-m\tau_0} + \frac{\lambda}{2} \int_0^1 A(m, \eta' \tau_0) \frac{d\eta'}{\eta'}\right] - \\ &- \psi(\eta, \tau_0) \left[1 + \frac{\lambda}{2} \int_0^1 B(m, \eta', \tau_0) \frac{d\eta'}{\eta'}\right]. \end{aligned} \tag{39}$$

We now introduce the notation

$$\begin{aligned} e^{-m\tau_0} + \frac{\lambda}{2} \int_0^1 A(m, \eta', \tau_0) \frac{d\eta'}{\eta'} &= a(m, \tau_0), \\ 1 + \frac{\lambda}{2} \int_0^1 B(m, \eta', \tau_0) \frac{d\eta'}{\eta'} &= b(m, \tau_0). \end{aligned} \tag{40}$$

Then, instead of (39) we find

$$A(m, \eta, \tau_0) = \frac{\eta}{1 - m\eta} [\varphi(\eta, \tau_0)\, a(m, \tau_0) - \psi(\eta, \tau_0)\, b(m, \tau_0)]. \tag{41}$$

Multiplying (29) by $e^{-m\tau}$ and integrating with respect to τ from 0 to τ_0, we get similarly

$$B(m, \eta, \tau_0) = \frac{\eta}{1 + m\eta} [\varphi(\eta, \tau_0)\, b(m, \tau_0) - \psi(\eta, \tau_0)\, a(m, \tau_0)]. \tag{42}$$

From (41) and (42) we obtain the following system of equations for determining the quantities $a(m, \tau_0)$ and $b(m, \tau_0)$:

$$\begin{aligned} a(m, \tau_0) &\left[1 - \frac{\lambda}{2} \int_0^1 \frac{\varphi(\eta, \tau_0)}{1 - m\eta} d\eta\right] + \\ &+ b(m, \tau_0) \frac{\lambda}{2} \int_0^1 \frac{\psi(\eta, \tau_0)}{1 - m\eta} d\eta = e^{-m\tau_0}. \end{aligned} \tag{43}$$

$$a(m, \tau_0) \frac{\lambda}{2} \int_0^1 \frac{\psi(\eta, \tau_0)}{1 + m\eta} d\eta + b(m, \tau_0) \left[1 - \frac{\lambda}{2} \int_0^1 \frac{\varphi(\eta, \tau_0)}{1 + m\eta} d\eta\right] = 1. \tag{44}$$

The determinant of the system of Equations (43) and (44) can be written in the form:

$$\Delta = 1 - \frac{\lambda}{2}\int\limits_0^1 \frac{\varphi(\eta,\tau_0)}{1-m\eta}\,d\eta - \frac{\lambda}{2}\int\limits_0^1 \frac{\varphi(\eta,\tau_0)}{1+m\eta}\,d\eta + $$

$$+\frac{\lambda^2}{4}\int\limits_0^1\int\limits_0^1 \frac{\varphi(\eta,\tau_0)\,\varphi(\zeta,\tau_0) - \psi(\eta,\tau_0)\,\psi(\zeta,\tau_0)}{(1-m\eta)(1+m\zeta)}\,d\eta\,d\zeta. \tag{45}$$

Because of (17), the last term in Equation (45) is equal to

$$\lambda\int\limits_0^1\int\limits_0^1 \frac{\rho(\eta,\zeta,\tau_0)(\eta+\zeta)}{(1-m\eta)(1+m\zeta)}\,d\eta\,d\zeta.$$

Since

$$\frac{\eta+\zeta}{(1-m\eta)(1+m\zeta)} = \frac{1}{m}\left(\frac{1}{1-m\eta} - \frac{1}{1+m\zeta}\right),$$

it can be rewritten in the following form:

$$\frac{\lambda}{m}\int\limits_0^1 \frac{d\eta}{1-m\eta}\int\limits_0^1 \rho(\eta,\zeta,\tau_0)\,d\zeta - \frac{\lambda}{m}\int\limits_0^1 \frac{d\eta}{1+m\eta}\int\limits_0^1 \rho(\eta,\zeta,\tau_0)\,d\zeta.$$

Consequently, we have

$$\Delta = 1 - \lambda\int\limits_0^1 \frac{\varphi(\eta,\tau_0)}{1-m^2\eta^2}\,d\eta + 2\lambda\int\limits_0^1 \frac{\eta\,d\eta}{1+m^2\eta^2}\int\limits_0^1 \rho(\eta,\zeta,\tau_0)\,d\zeta, \tag{46}$$

or, taking (10) into account,

$$\Delta = 1 - \lambda\int\limits_0^1 \frac{d\eta}{1-m^2\eta^2} = 1 - \frac{\lambda}{2m}\log\frac{1+m}{1-m}\,. \tag{47}$$

If m = k, Δ = 0. Therefore, Equations (43) and (44) yield the following two relations:

$$\frac{\lambda}{2}\int\limits_0^1 \frac{\varphi(\eta,\tau_0)}{1-k\eta}\,d\eta + e^{-k\tau_0}\,\frac{\lambda}{2}\int\limits_0^1 \frac{\psi(\eta,\tau_0)}{1+k\eta}\,d\eta = 1; \tag{48}$$

$$\frac{\lambda}{2}\int\limits_0^1 \frac{\varphi(\eta,\tau_0)}{1+k\eta}\,d\eta + e^{k\tau_0}\,\frac{\lambda}{2}\int\limits_0^1 \frac{\psi(\eta,\tau_0)}{1-k\eta}\,d\eta = 1. \tag{49}$$

Relations (48) and (49) can be used to find the constants C and C_1. Substituting (36) and (37) into (48) and (49), we find

$$C\int\limits_0^1 \frac{\varphi(\eta)}{(1-k\eta)^2}\,\eta\,d\eta = C_1\int\limits_0^1 \frac{\varphi(\eta)}{1-k^2\eta^2}\,\eta\,d\eta; \tag{50}$$

$$C_1\int\limits_0^1 \frac{\varphi(\eta)}{(1-k\eta)^2}\,\eta\,d\eta = 2k\int\limits_0^1 \frac{\varphi(\eta)}{1-k^2\eta^2}\,\eta\,d\eta. \tag{51}$$

From (50) and (51), incidentally, it follows that

$$2kC = C_1^2. \tag{52}$$

It should be noted that the asymptotic expressions (36) and (37) for the functions $\varphi(\eta, \tau_0)$ and $\psi(\eta, \tau_0)$ could have been obtained also from Equations (19), (20) and (23), (24). We would also have obtained Relation (52) for the two constants C and C_1. However, to determine the constant C_1 it would be necessary to derive Equation (51) again.

The case when $\lambda = 1$. In this case k = 0 and Expressions (36) and (37) cannot be used. From the equations given above which determine the functions $\varphi(\eta, \tau_0)$ and $\psi(\eta, \tau_0)$ we find for $\lambda = 1$ and $\tau \gg 1$

$$\varphi(\eta, \tau_0) = \varphi(\eta) - \frac{\eta\varphi(\eta)}{\tau_0 + \gamma}; \tag{53}$$

$$\psi(\eta, \tau_0) = \frac{\eta\varphi(\eta)}{\tau_0 + \gamma}, \tag{54}$$

where γ is a constant.

To determine γ we make use of the fact that in the case of pure scattering the probability for the escape of a quantum from any depth in any direction is equal to unity, i. e.,

$$2\pi\int_0^1 p(\tau, \eta, \tau_0)\, d\eta + 2\pi\int_0^1 p(\tau_0 - \tau, \eta, \tau_0)\, d\eta = 1. \tag{55}$$

When $\tau = 0$, from (55) it follows that

$$\int_0^1 \varphi(\eta, \tau_0)\, d\eta + \int_0^1 \psi(\eta, \tau_0)\, d\eta = 2. \tag{56}$$

However, Equation (56) cannot be used for finding γ, since γ cancels out when (53) and (54) are substituted into (56).

In order to find γ, we integrate (55) with respect to τ from 0 to τ_0. Using the notation

$$A(0, \eta, \tau_0) = 4\pi\int_0^{\tau_0} p(\tau, \eta, \tau_0)\, d\tau, \tag{57}$$

we get

$$\int_0^1 A(0, \eta, \tau_0)\, d\eta = \tau_0. \tag{58}$$

On the other hand, the quantity $A(0, \eta, \tau_0)$ can be found from Equation (29). Putting m = 0 and $\lambda = 1$ in Equations (41) and (42), which were obtained from (29), we have

$$A(0, \eta, \tau_0) = \eta \frac{\varphi(\eta, \tau_0) - \psi(\eta, \tau_0)}{1 - \frac{1}{2}\int_0^1 [\varphi(\eta', \tau_0) - \psi(\eta', \tau_0)]\, d\eta'}. \tag{59}$$

When $\tau_0 \gg 1$, Equation (59) with the help of (53) and (54) gives

$$A(0, \eta, \tau_0) = \eta\varphi(\eta)\frac{\tau_0 + \gamma - 2\eta}{\alpha_1}, \tag{60}$$

where we have used the notation

$$\alpha_k = \int_0^1 \varphi(\eta)\,\eta^k\,d\eta. \tag{61}$$

Substituting (60) into (58), we finally obtain the required expression for γ:

$$\gamma = 2\frac{\alpha_2}{\alpha_1}. \tag{62}$$

3. Emission from a Medium with Various Radiation Sources

If the function $p(\tau, \eta, \tau_0)$ is known for a given medium, then the intensity of the radiation emerging from the medium with any radiation sources can be obtained by means of a suitable integration. Let $f(\tau)\,d\tau$ denote the amount of energy directly supplied by the radiation sources and absorbed per second by an element of volume at a depth τ whose cross sectional area is 1 cm^2 and optical thickness is $d\tau$. Then, the intensity of radiation emerging from the medium through the upper ($\tau = 0$) and lower boundaries ($\tau = \tau_0$) will be equal to

$$I(0, \eta, \tau_0) = \int_0^{\tau_0} p(\tau, \eta, \tau_0)\,f(\tau)\,\frac{d\tau}{\eta}; \tag{63}$$

$$I(\tau_0, \eta, \tau_0) = \int_0^{\tau_0} p(\tau_0 - \tau, \eta, \tau_0)\,f(\tau)\,\frac{d\tau}{\eta}. \tag{64}$$

Let us assume that the radiation sources are situated inside the medium and that they emit radiation uniformly in all directions. Let $g(\tau)\,d\tau$ denote the amount of energy radiated per second into a unit solid angle by the radiation sources situated at an optical depth τ in an element of volume 1 cm^2 in cross section and of optical thickness $d\tau$. It is obvious that we can take

$$f(\tau) = \frac{4\pi}{\lambda}\,g(\tau). \tag{65}$$

Therefore, instead of (63) and (64) we have

$$I(0, \eta, \tau_0) = \frac{4\pi}{\lambda}\int_0^{\tau_0} p(\tau, \eta, \tau_0)\,g(\tau)\,\frac{d\tau}{\eta}; \tag{66}$$

$$I(\tau_0, \eta, \tau_0) = \frac{4\pi}{\lambda}\int_0^{\tau_0} p(\tau_0 - \tau, \eta, \tau_0)\,g(\tau)\,\frac{d\tau}{\eta}. \tag{67}$$

In many cases the quantities $I(0, \eta, \tau_0)$ and $I(\tau_0, \eta, \tau_0)$ can be found without initially determining the function $p(\tau, \eta, \tau_0)$ directly from Equation (29). This leads to an expression for the intensity of the radiation emerging from the medium in terms of the functions $\varphi(\eta, \tau_0)$ and $\psi(\eta, \tau_0)$. For a medium with a very large optical thickness, $I(0, \eta, \tau_0)$ and $I(\tau_0, \eta, \tau_0)$ can be expressed in terms of the function $\varphi(\eta)$ with the help of the equations derived in the previous section.

Examples showing the method for determining the intensity of the radiation emerging from a medium with various arrangements of the radiation sources are given below.

1) Let us assume that parallel rays, which on an area normal to them produce an illumination πS, fall on the upper boundary of the medium at an angle arc cos ζ to the normal. In this case

$$f(\tau) = \pi S e^{-\tau/\zeta}, \tag{68}$$

and the intensities of the radiation diffusely scattered and diffusely transmitted by the medium, from Equations (63) and (64), are given by

$$I(0, \eta, \tau_0) = \pi S\int_0^{\tau_0} p(\tau, \eta, \tau_0)\,e^{-\tau/\zeta}\,\frac{d\tau}{\eta}; \tag{69}$$

$$I(\tau_0, \eta, \tau_0) = \pi S \int\limits_0^{\tau_0} p(\tau_0 - \tau, \eta, \tau_0)\, e^{-\tau/\zeta}\, \frac{d\tau}{\eta}. \tag{70}$$

Comparing Equations (69), (70) with (7), (8), we have

$$I\ (0, \eta, \tau_0) = S\rho(\eta, \zeta, \tau_0)\,\zeta,\ \ I(\tau_0, \eta, \tau_0) = S\sigma(\eta, \zeta, \tau_0)\,\zeta. \tag{71}$$

The quantities $I(0, \eta, \tau_0)$ and $I(\tau_0, \eta, \tau_0)$ can be expressed in terms of the functions $\varphi(\eta, \zeta, \tau_0)$ and $\sigma(\eta, \zeta, \tau_0)$, which were introduced earlier, by means of Equations (71). Moreover, from Equations (71) we can obtain the physical meaning of the coefficients of brightness $\rho(\eta, \zeta, \tau_0)$ and $\sigma(\eta, \zeta, \tau_0)$: the coefficient of brightness is the ratio of the intensity of radiation, diffusely reflected or diffusely transmitted by the medium, to the intensity of radiation diffusely reflected from an orthotropic ideally-white area situated on the upper boundary of the medium.

The coefficients of brightness can be expressed in terms of the functions $\varphi(\eta, \tau_0)$ and $\psi(\eta, \tau_0)$ by means of Equations (17), (18) or (21), (22) or (25), (26). Substituting (36) and (37) into the equations indicated above, we obtain the following asymptotic expressions for the coefficients of brightness when $\lambda < 1$:

$$\rho(\eta, \zeta, \tau_0) = \frac{\lambda}{4} \frac{\varphi(\eta)\varphi(\zeta)}{\eta + \zeta} - \frac{\lambda}{4} C \frac{\varphi(\eta)}{1 - k\eta} \frac{\varphi(\zeta)}{1 - k\zeta} e^{-2k\tau_0}, \tag{72}$$

$$\sigma(\eta, \zeta, \tau_0) = \frac{\lambda}{4} C_1 \frac{\varphi(\eta)}{1 - k\eta} \frac{\varphi(\zeta)}{1 - k\eta} e^{-k\tau_0}. \tag{73}$$

Similarly, when $\lambda = 1$ we find with the help of Equations (53) and (54) that

$$\rho(\eta, \zeta, \tau_0) = \frac{1}{4} \frac{\varphi(\eta)\varphi(\zeta)}{\eta + \zeta} - \frac{1}{4} \frac{\varphi(\eta)\varphi(\zeta)}{\tau_0 + \gamma}; \tag{74}$$

$$\sigma(\eta, \zeta, \tau_0) = \frac{1}{4} \frac{\varphi(\eta)\varphi(\zeta)}{\tau_0 + \gamma}. \tag{75}$$

Equations (72), (73), (74) and (75) were earlier obtained by V. A. Ambartsumian by a different method (see Reference [2], Chapter III).

2) We will assume that

$$g(\tau) = Ge^{-m\tau}, \tag{76}$$

where G and $\underline{m}$ are constants. Substituting (76) into Equations (66) and (67) and then comparing the resulting expressions with (38), we find

$$I(0, \eta, \tau_0) = \frac{G}{\eta} B(m, \eta, \tau_0), \quad I(\tau_0, \eta, \tau_0) = \frac{G}{\eta} A(m, \eta, \tau_0), \tag{77}$$

or, taking (41) and (42) into account,

$$I(0, \eta, \tau_0) = \frac{G}{1 + m\eta} [\varphi(\eta, \tau_0)\, b(m, \tau_0) - \psi(\eta, \tau_0)\, a(m, \tau_0)]; \tag{78}$$

$$I(\tau_0, \eta, \tau_0) = \frac{G}{1 - m\eta} [\varphi(\eta, \tau_0)\, a(m, \tau_0) - \psi(\eta, \tau_0)\, b(m, \tau_0)]. \tag{79}$$

The quantities $a(m, \tau_0)$ and $b(m, \tau_0)$ appearing in (78) and (79) are determined from Equations (43) and (44). We have

$$a(m, \tau_0) = \frac{1}{\Delta(m)} \left\{e^{-m\tau_0} - e^{-m\tau_0} \frac{\lambda}{2} \int\limits_0^1 \frac{\varphi(\eta, \tau_0)}{1 + m\eta} d\eta - \frac{\lambda}{2} \int\limits_0^1 \frac{\psi(\eta, \tau_0)}{1 - m\eta} d\eta\right\}; \tag{80}$$

$$b(m, \tau_0) = \frac{1}{\Delta(m)} \left\{1 - \frac{\lambda}{2} \int_0^1 \frac{\varphi(\eta, \tau_0)}{1 - m\eta}\, d\eta - e^{-m\tau_0} \frac{\lambda}{2} \int_0^1 \frac{\psi(\eta, \tau_0)}{1 + m\eta}\, d\eta\right\}, \tag{81}$$

where $\Delta(m)$ is given by Equation (47).

It must be noted that when $m = \frac{1}{\zeta} \geq 1$ and $G = \frac{\lambda}{4} S$, Equations (78) and (79) transform into Equations (17) and (18), since the following relations hold:

$$a\left(\frac{1}{\zeta}, \tau_0\right) = \psi(\zeta, \tau_0), \quad b\left(\frac{1}{\zeta}, \tau_0\right) = \varphi(\zeta, \tau_0), \tag{82}$$

which can be obtained from a comparison of Equations (38) and (40) with Equations (7), (8) and (10), (11).

Using (82), we can write Equations (80) and (81) as:

$$\varphi(\zeta, \tau_0)\, \Delta\left(\frac{1}{\zeta}\right) = 1 + \frac{\lambda}{2} \zeta \int_0^1 \frac{\varphi(\eta, \tau_0)}{\eta - \zeta}\, d\zeta - e^{-\tau_0/\zeta} \frac{\lambda}{2} \zeta \int_0^1 \frac{\psi(\eta, \tau_0)}{\eta + \zeta}\, d\eta; \tag{83}$$

$$\psi(\zeta, \tau_0)\, \Delta\left(\frac{1}{\zeta}\right) = e^{-\tau_0/\zeta} - e^{-\tau_0/\zeta} \frac{\lambda}{2} \zeta \int_0^1 \frac{\varphi(\eta, \tau_0)}{\eta + \zeta}\, d\eta + \frac{\lambda}{2} \zeta \int_0^1 \frac{\psi(\eta, \tau_0)}{\eta - \zeta}\, d\eta. \tag{84}$$

The author has obtained Equations (83) and (84) earlier by another method [7].

When $m = k$, the numerator and denominator of Equations (80) and (81) become zero. Eliminating this indeterminacy, we get

$$a(k, \tau_0) = \frac{\lambda}{2\Delta'(k)} \left\{e^{-k\tau_0} \int_0^1 \frac{\varphi(\eta, \tau_0)}{(1 + k\eta)^2}\, \eta\, d\eta - \tau_0 \int_0^1 \frac{\psi(\eta, \tau_0)}{1 - k\eta}\, d\eta - \int_0^1 \frac{\psi(\eta, \tau_0)}{(1 - k\eta)^2}\, \eta\, d\eta\right\}, \tag{85}$$

$$\begin{aligned} b(k, \tau_0) = \frac{\lambda}{2\Delta'(k)} \Bigg\{ &- \int_0^1 \frac{\varphi(\eta, \tau_0)}{(1 - k\eta)^2}\, \eta\, d\eta + \\ &+ \tau_0 e^{-k\tau_0} \int_0^1 \frac{\psi(\eta, \tau_0)}{1 + k\eta}\, d\eta + e^{-k\tau_0} \int_0^1 \frac{\psi(\eta, \tau_0)}{(1 + k\eta)^2}\, \eta\, d\eta \Bigg\}, \end{aligned} \tag{86}$$

where

$$\Delta'(k) = \frac{1}{k} \left(1 - \frac{\lambda}{1 - k^2}\right). \tag{87}$$

Let us find the asymptotic expressions for the quantities $I(0, \eta, \tau_0)$ and $I(\tau_0, \eta, \tau_0)$ for the case of large optical thickness.

When $\lambda < 1$, the substitution of (36) and (37) into (80) and (81) yields

$$a(m, \tau_0) = \frac{\Phi(m)}{\Delta(m)} \left(1 + \frac{C}{k + m} e^{-2k\tau_0}\right) e^{-m\tau_0} - \frac{\Phi(-m)}{\Delta(m)} \frac{C_1}{k - m} e^{-k\tau_0}; \tag{88}$$

$$b(m, \tau_0) = \frac{\Phi(-m)}{\Delta(m)} \left(1 + \frac{C}{k - m} e^{-2k\tau_0}\right) - \frac{\Phi(m)}{\Delta(m)} \frac{C_1}{k + m} e^{-(m+k)\tau_0}, \tag{89}$$

where we have introduced the notation

$$\Phi(m) = 1 - \frac{\lambda}{2} \int_0^1 \frac{\varphi(\eta)}{1 + m\eta}\, d\eta. \tag{90}$$

When $m < k$, substituting (88), (89) into (78), (79) and neglecting all but the largest terms, we find

$$I(0, \eta, \tau_0) = \frac{G\varphi(\eta)}{1 + m\eta} \frac{\Phi(-m)}{\Delta(m)} - \frac{G\varphi(\eta)}{(1 - k\eta)(k + m)} \frac{\Phi(m)}{\Delta(m)} C_1 e^{-(k+m)\tau_0}; \tag{91}$$

$$I(\tau_0, \eta, \tau_0) = \frac{G\varphi(\eta)}{1 - m\eta} \frac{\Phi(m)}{\Delta(m)} e^{-m\tau_0}. \tag{92}$$

Similarly, when $m > k$ we get

$$I(0, \eta, \tau_0) = G\varphi(\eta) \frac{\Phi(-m)}{\Delta(m)} \left[\frac{1}{1 + m\eta} - \frac{C}{(1 - k\eta)(m - k)} e^{-2k\tau_0} \right], \tag{93}$$

$$I(\tau_0, \eta, \tau_0) = \frac{G\varphi(\eta)}{(1 - k\eta)(m - k)} \frac{\Phi(-m)}{\Delta(m)} C_1 e^{-k\tau_0}. \tag{94}$$

It should be noted that with $m = \dfrac{1}{\zeta} \geq 1$ and $G = \dfrac{\lambda}{4} S$, Equations (93) and (94) transform into Equations (72) and (73) because of the existence of the relation

$$\varphi(\zeta)\,\Delta\left(\frac{1}{\zeta}\right) = 1 + \frac{\lambda}{2} \zeta \int_0^1 \frac{\varphi(\eta)}{\eta - \zeta} d\eta, \tag{95}$$

which is obtained from (83) as $\tau_0 \rightarrow \infty$.

If $m = k$, then to find the quantities $I(0, \eta, \tau_0)$ and $I(\tau_0, \eta, \tau_0)$ we must substitute the Expressions (85) and (86) into the Equations (78) and (79). For $\tau_0 >> 1$ we have

$$I(0, \eta, \tau_0) = \frac{G}{1 + k\eta} \frac{\varphi(\eta)}{\Delta'(k)} \frac{\lambda}{2} \left\{ -\int_0^1 \frac{\varphi(\eta')}{(1 - k\eta')^2} \eta' d\eta' + \tau_0 e^{-2k\tau_0} C_1 \int_0^1 \frac{\varphi(\eta')}{1 - k^2\eta'^2} \eta d\eta \right\}; \tag{96}$$

$$I(\tau_0, \eta, \tau_0) = -\frac{G}{1 - k\eta} \frac{\varphi(\eta)}{\Delta'(k)} \tau_0 e^{-k\tau_0} \lambda k \int_0^1 \frac{\varphi(\eta')}{1 - k^2\eta'^2} \eta' d\eta'. \tag{97}$$

When $\lambda = 1$, we must use the Expressions (53) and (54) for the functions $\varphi(\eta, \tau_0)$ and $\psi(\eta, \tau_0)$. Substituting them into Equations (80) and (81) and (78), (79), we find

$$I(0, \eta, \tau_0) = \frac{G\varphi(\eta)}{m\Delta(m)} \left[1 - \frac{1}{2} \int_0^1 \frac{\varphi(\eta')}{1 - m\eta'} d\eta' \right] \left(\frac{m}{1 + m\eta} - \frac{1}{\tau_0 + \gamma} \right); \tag{98}$$

$$I(\tau_0, \eta, \tau_0) = \frac{G\varphi(\eta)}{m\Delta(m)} \left[1 - \frac{1}{2} \int_0^1 \frac{\varphi(\eta')}{1 - m\eta'} d\eta' \right] \frac{1}{\tau_0 + \gamma}. \tag{99}$$

It can be easily seen that when $m = \dfrac{1}{\zeta} \geq 1$ and $G = \dfrac{S}{4}$, Equations (98) and (99) transform into Equations (74) and (75).

When $\lambda = 1$, all of the energy produced inside the medium escapes from it. If $g(\tau) = Ge^{-m\tau}$ and $\tau_0 >> 1$, then we must have

$$\int_0^1 [I(0, \eta, \tau_0) + I(\tau_0, \eta, \tau_0)]\, \eta d\eta = \frac{2}{m} G. \tag{100}$$

Substituting (98) and (99) into (100), we get a particular case (for $\lambda = 1$) of the following relation:

$$\Phi(m)\,\Phi(-m) = \Delta(m), \tag{101}$$

which can be obtained from Equation (30).

3) Let us assume that the energy sources are distributed uniformly throughout the medium, i. e., $g(\tau) = G = \text{const}$. In this case

$$I(0, \eta, \tau_0) = I(\tau_0, \eta, \tau_0) = \frac{4\pi}{\lambda} G \int_0^1 p(\tau, \eta, \tau_0) \frac{d\tau}{\eta}. \tag{102}$$

The integral appearing in (102) can be found from Equation (29). However, we can also use the Relations (38), (41), and (43) for this purpose. Taking m = 0 in them, we get

$$I(0, \eta, \tau_0) = G \frac{\varphi(\eta, \tau_0) - \psi(\eta, \tau_0)}{1 - \frac{\lambda}{2}\int_0^1 \varphi(\eta, \tau_0)\, d\eta + \frac{\lambda}{2}\int_0^1 \psi(\eta, \tau_0)\, d\eta}. \tag{103}$$

For $\tau_0 \gg 1$ and $\lambda < 1$ we find, after substituting Equations (36) and (37) into (103),

$$I(0, \eta, \tau_0) = G\varphi(\eta) \frac{1 - \frac{\eta}{1 - k\eta} C_1 e^{-k\tau_0}}{1 - \frac{\lambda}{2}\int_0^1 \varphi(\eta)\, d\eta + C_1 e^{-k\tau_0} \frac{\lambda}{2}\int_0^1 \frac{\varphi(\eta)}{1 - k\eta}\eta\, d\eta}. \tag{104}$$

But, from (30) it follows that

$$1 - \frac{\lambda}{2}\int_0^1 \varphi(\eta)\, d\eta = \sqrt{1 - \lambda}, \tag{105}$$

while

$$\frac{\lambda}{2}\int_0^1 \frac{\varphi(\eta)}{1 - k\eta}\eta\, d\eta = \frac{\lambda}{2k}\left[\int_0^1 \frac{\varphi(\eta)}{1 - k\eta}\, d\eta - \int_0^1 \varphi(\eta)\, d\eta\right] = \frac{\sqrt{1 - \lambda}}{k}. \tag{106}$$

Therefore, instead of (104) we get

$$I(0, \eta, \tau_0) = G\frac{\varphi(\eta)}{\sqrt{1 - \lambda}}\left(1 - \frac{C_1}{k} e^{-k\tau_0} \frac{1}{1 - k\eta}\right). \tag{107}$$

It should be noted that Equation (107) can also be obtained from Equation (91) by putting m = 0 (since $\Delta(0) = 1 - \lambda$).

When $\tau_0 \gg 1$ and $\lambda = 1$, substituting Expressions (53) and (54) into (103), we get

$$I(0, \eta, \tau_0) = G\varphi(\eta)\, \frac{\tau_0 + \gamma - 2\eta}{\alpha_1}. \tag{108}$$

With the help of Equation (29) we can also obtain expressions for the intensity of the radiation escaping from the medium in the case when $g(\tau)$ is a polynomial in τ. These expressions were obtained in Chapter VI, Section 7 of our book [2]. Substituting the expressions for the functions $\varphi(\eta, \tau_0)$ and $\psi(\eta, \tau_0)$ which were found above into these equations, we can obtain the asymptotic expressions for the quantities $I(0, \eta, \tau_0)$ and $I(\tau_0, \eta, \tau_0)$ for large values of τ_0. However, we will not carry out this calculation here.

A. A. Zhdanov Leningrad State University

Received December 6, 1956

LITERATURE CITED

[1] V. V. Sobolev, Astron. J. (USSR) 28, 5, 335 (1951).

[2] V. V. Sobolev, The Transport of Radiant Energy in the Atmospheres of the Stars and Planets [in Russian] (Gostekhizdat, 1956).

[3] V. A. Ambartsumian, Doklady Akad. Nauk SSSR 37, 8, 257 (1943).

[4] V. V . Sobolev, Doklady Akad. Nauk SSSR 111, 5 (1956). *

[5] S. Chandrasekhar, Donna Elbert and Ann Franklin, Astrophys. J. 115, 2, 244 (1952).

[6] J. W. Busbridge, Astrophys. J. 122, 2, 327 (1955).

[7] V. V. Sobolev, Doklady Akad. Nauk SSSR 69, 4 (1949).

---

*[Soviet Physics — Doklady, p. 1000.]

RADIO STARS AND THE METHODS OF THEIR INVESTIGATION*

V. V. Vitkevich

In § 1 Ryle's results on the spatial distribution of radio stars are discussed.

In § 2 the experimental curve, derived by Ryle, which gives the dependence of N on I (I is the intensity of radio stars, N is the number of radio stars per one steradian) is used for calculating the function giving the variation of the density of the sources $\rho(r)$ with distance, assuming that their luminosity is constant, and also the variation of luminosity $L(r)$, assuming that the density is constant.

In order to choose between these possibilities, the problem of the measurement of angular dimensions of radio stars as a function of their intensity and number is discussed. It is shown that such observations are possible and that it will be possible to decide in favor of one of the hypotheses.

Further (§ 3), the question of the form of radio-reception diagrams of large radio telescopes is considered. It is shown that "elongated" diagrams with a large resolving power only in one direction are not suitable for the further investigation of the above mentioned problem. The "elongated" form of a diagram does not permit the same resolving power to be obtained along the second coordinate as along the first.

In § 4 general relations between the sensitivity and resolving power of large telescopes in connection with the problem under discussion, are considered. It is shown that the resolving power of large radio telescopes of the ordinary type (parabolic reflectors, synphase antennas, etc) is not sufficiently high. Their sensitivity increases more rapidly with increasing area in the range of metric waves than the resolving power and they are of little use for the further investigation of the problem under discussion.

Antenna systems of the "cross" type which have a much higher resolving power, are the most suitable for this investigation.

In § 5 attention is drawn to the fact that side lobes of large antenna systems can be a serious hindrance when searching for the largest number of radio stars.

In § 6 the supposition that the future of large radio telescopes belongs to radioreflectors and not radio-refractors is criticized. It is shown that radio-refractors have larger possibilities than radio-reflectors.

§ 1. Ryle's Work on the Observation of Radio Stars

In recent years, because of the large number of radio stars discovered, the radioastronomers have been devoting more and more attention to the investigation of radio stars.

---

*Presented at the symposium of the Crimean Station of the FIAN on June 21, 1956.

New observational data of fundamental scientific interest have been published recently. We refer to the work carried out under Ryle's supervision [1, 2], whose results we will discuss here.

An interferometer consisting of four parabolic cylinders of size 320 ft × 40 ft and with a total area of about 50,000 sq. ft. was used for the investigations. The measurements were made at a wavelength of 3.7 m.

With the help of this interferometer 1936 radio stars were discovered and the position of 500 of these was determined with a very high accuracy — 2 minutes of arc in right ascension and 12 minutes of arc in declination.

Thirty of these sources have fairly large angular dimensions, from 20' to 180', and the more intense sources were found to be situated within 5° of the galactic equator. One of them was identified with the nebula IC443 and according to the author, the other extended sources are of the same type. The remaining 1906 sources have small angular dimensions.

The study of the positions of the 75 most powerful sources carried out by Minkowski, has shown that only a few of them can be identified with faint galaxies.

It was established that probably about 80 radio stars correspond to objects listed in the NGC catalogue and the Index Cat., some of which are double or multiple galaxies. There are reasons for believing that one of the sources corresponds to NGC 1275 and Baade concluded that this is a second example of colliding galaxies.

However, even if we consider that in future some of the other sources will be identified, nevertheless the identified sources will only be a small fraction of the total number of observed radio stars.

As regards the remaining radio stars, it was found that they are distributed isotropically in space. A special exploration was made of 25 small areas of the celestial sphere and it was found that there is no observed variation in the distribution of radio stars.

The frequency distribution of the radio-star intensities was next obtained from the data.

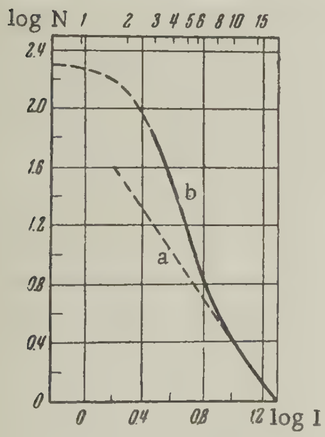


Fig. 1. Log N as a function of log I (N is the number of sources per steradian, I is the intensity), a) the theoretical curve for a uniform spatial distribution of the sources; b) Ryle's experimental curve.

Let N denote the number of radio stars within a solid angle equal to one steradian and having an intensity equal to or greater than I. Then the observed frequency distribution of I (on a logarithmic scale) is given by the curve b in Figure 1.

In order to discuss this result, we will consider a model in which the sources are distributed isotropically in space.

If the brightness of the source is L, then its intensity at a distance r will be $I = Lr^{-2}$; $r = L^{1/2} I^{-1/2}$.

If the sources are uniformly distributed in space with a density ρ, then the number of them situated within a unit solid angle up to a distance r is

$$N = \int_0^r \rho r^2 \, dr = \frac{1}{3} \rho r^3. \tag{1}$$

Assuming that all the sources have the same value of L, we find that the number of sources with an intensity equal to or greater than I is given by

$$N = \frac{1}{3} \rho L^{3/2} I^{-3/2}; \quad \log N = -\frac{3}{2} \log I + \text{const.} \tag{2}$$

Thus, the slope of the log N — log I curve must be −3/2. This is the case for isotropically-distribued sources (Fig. 1,a). Calculations show that this equation holds also when there is a distribution in the luminosity L and the density ρ. In this case L and ρ must be understood as mean values.

As can be seen from a comparison of the curves (Fig. 1,a and 1, b), at first for small values of N the experimental curve is fairly close to the theoretical, which indicates that the sources are distributed isotropically; however, for higher values of N, the experimental curve rises sharply. The observed rise of the curve cannot, apparently, be associated with any experimental effects; it is either caused by an increase of the source density ρ or an increase of their luminosity L with distance from the solar system. This result is of fundamental importance and requires a thorough verification and study.

§ 2. Spatial Distribution of Radio Stars. Angular Dimensions

In order to discuss the problem indicated and to map out a plan for the future solution of the problem, which is one of the main problems in modern radioastronomy and cosmology, we will derive some relations which might prove useful and also examine in detail some of the consequences of curve b in Figure 1.

Firstly, let us investigate what type of relation between I and N we can expect if the luminosity L and the source density ρ vary with the distance.

The general equations for N and I are

$$N(r) = \int_0^r \rho(r) r^2 \, dr; \quad I = L(r) r^{-2}. \tag{3}$$

The elimination of r leads to an expression for N as a function of I. On the other hand, knowing N as a function of I from experiment, we can deduce the dependence of N and L on r if we assume that one of these functions is known (for example, independent of r).

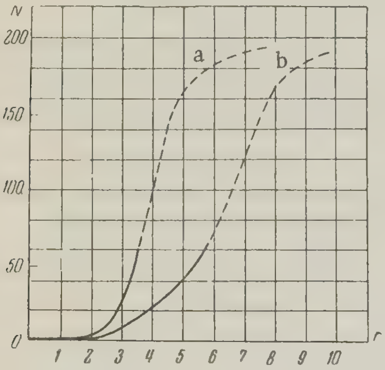


Fig. 2. The number of radio stars N (per steradian) as a function of the distance: a) $\rho \neq$ const; L = const; b) ρ = const; L $\neq$ const.

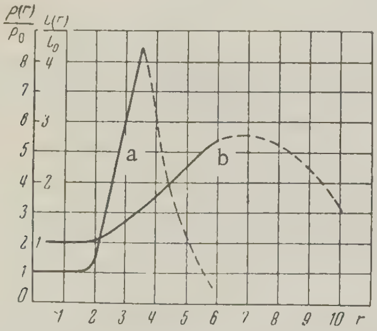


Fig. 3. a) The variation of the density of radio stars ρ with distance on the assumption that L = const; b) the variation of the luminosity of radio stars as a function of the distance on the assumption that ρ = const.

Let us determine the spatial distribution of the sources (density ρ) or of the luminosity (L), assuming that the data given in Figure 1 are correct.

For the region in which the variation of N with I is governed by the exponent n = 3/2, we can obtain the following initial values of all the quantities (assuming that the density equals unity):

$$\rho = 1; \; N_0 = 2.5; \; I_0 = 10.7; \; r_0 = 2; \; L_0 = 40. \tag{4}$$

These values are used as boundary conditions for the investigation of the following two cases:

1) $L = L_0$ = const. We then have

$$I = L_0 r^{-2}; \; \log I = -2 \log r + \lg L_0. \tag{5}$$

Let us find r from the obvious relation:

$$r = L_0^{1/2} I^{-1/2}$$

Next, we obtain the function N = N(r) (Fig. 2, a).

In order to obtain the function $\rho(r)$, we use the following obvious equations:

$$N = N_0 + \int_{r_0}^{r} \rho(r) r^2 \, dr;$$

$$\Delta N(r_2, r_1) = \int_{r_1}^{r_2} \rho(r) r^2 \, dr \approx \rho\left(\frac{r_1 + r_2}{2}\right) \frac{r_2^3 - r_1^3}{3}, \tag{6}$$

where ΔN is the variation in the number of observed radio stars when the distance is varied from r_1 to r_2.

Knowing the consecutive values of N for close values of r and, consequently, knowing $\Delta N(r_2, r_1)$, we find the values of ρ. The results of this calculation are given in Figure 3, a. As can be seen, ρ is constant for $r < 2$, while for larger r the density increases sharply, reaches a maximum equal to 8.5 and then falls.

Here, and in the other figures, the dashed line indicates the section of the curve which is in the region of "saturation", where the maximum number of observed sources is limited by the insufficient resolving power of the apparatus; this point will be discussed in greater detail in Section 4.

We will now consider the other possibility.

2) $\rho = \rho_0 = \text{const}$; $L = L(r)$.

First of all we can find the distance r, which in this case is obviously given by

$$r = [3N(r)]^{1/3} \rho^{-1/3},$$

so that

$$L(r) = Ir^2 = \left[\frac{3N(r)}{\rho}\right]^{2/3}, \tag{7}$$

assuming $\rho = \rho_0 = 1$, we find the function L(r). The results of the calculation are given graphically in Figure 2, b and since it is immaterial what the initial value of L_0 is, relative values are given, i. e., $L(r)/L_0$.

As can be seen, the value of L at first remains constant, then it increases reaching the value 2.75 and finally decreases because of the "saturation".

Since in the second case the scale of distances r is different from that in the first, the function N(r) will have another form as can be seen from Figure 2, b.

At this stage it is reasonable to inquire: which of these two possibilities does in fact occur in nature? The results obtained and the method of investigation are not adequate for us to be able to answer this question.

Nevertheless, in principle, we can answer this question if we widen the scope of the observations.

With this aim, let us discuss what we can obtain from a measurement of the sizes of the observed radio stars.

If φ_0 is the angular size of a source at a distance r_0, then assuming that, on the average, the sizes of the sources are equal, we find

$$\varphi = \varphi_0 r_0 r^{-1}. \tag{8}$$

Using the relation $I = Lr^{-2}$, we get

$$\varphi = I^{1/2} L^{-1/2} \varphi_0 r_0; \tag{9}$$

Further, assuming that in general

$$N = N_0 I_0^n I^{-n}, \tag{10}$$

we can obtain the following relation between the angle φ and the number of sources N:

$$\varphi = N_0^{1/2} I_0^{1/2} N^{-1/2n} L^{-1/2} \varphi_0 r_0. \tag{11}$$

These are the general relations.

We can find the values of φ for the observed case given by the curve in Figure 1; in doing this we must again consider two possibilities: a) $L = \text{const}$, $\rho \neq \text{const}$ and, conversely, b) $\rho = \text{const}$, $L \neq \text{const}$.

For the first case, using the curve in Figure 2, a, we get the curve a in Figure 4. For the second case, using the curve in Figure 2, b, we get the curve b in Figure 4. For convenience we have taken $\varphi_0 = 1$ when $r = 2$.

As can be seen, the two curves deviate markedly from each other. For the case L = const, the distances r are found to be less, the angles φ correspondingly greater, than those for the case ρ = const.

For values of N between 20 and 60, i. e., when there is still no "saturation", the ratio of the values of φ for the two cases is equal to 2 which, undoubtedly, could be measured reliably.

We now note the following important point. In the preceding discussion it was tacitly assumed that the sizes of the radio stars are on the average constant and do not vary with distance. Let us assume that as the luminosity of the radio stars increases, the size changes in such a manner that their brightness remains constant. It is obvious that in this case the measurement of the angular dimensions will give a distribution similar to curve a in Figure 4. Consequently, by investigating the distribution of the angular dimensions of radio stars, we will be able to decide between the following two alternative hypotheses: 1) that the luminosity changes and 2) that the density changes or the radio-star size changes, the brightness remaining constant.

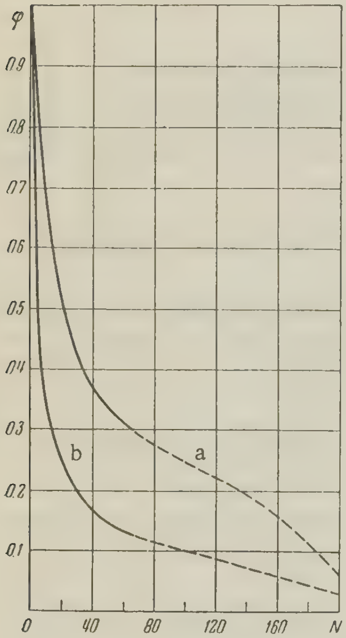


Fig. 4. Calculated values of the angular dimensions of radio stars: a) The case L = const, b) the case ρ = const.

If as the result of measurements it is found that the second alternative is to be preferred, then we must attempt to separate the case of constant density from the case of constant dimensions by comparing the spectra of the nearer sources with the spectra of the more distant ones. If the sizes of the radio stars change (their density remaining constant), they will be farther removed than if their sizes were constant. Therefore, in the first case the effect of the red shift and of attenuation will be more pronounced in the first case then in the second. Consequently, the differences in the spectra of the near and distant radio stars will also be appreciably greater in the first case than in the second. It is desirable to investigate these possibilities in greater detail after more experimental data have been obtained.

Turning now to the problem of the feasibility of measuring the size of radio stars, we note that the measurements could be carried out with an interferometer having the same type of antenna as was used by Ryle. However, this would not be an easy task since it would be necessary to place the antennas at very large distances from each other.

In fact, if we assume that $\varphi_0 = 1'$, then for a reliable measurement of the size of sources situated at a distance three times as great, a resolving power of the order of 0.3' would be required; the distance between the antennas would have to be $10^4 \lambda$ which, at a wavelength $\lambda = 3.7$ m, is about 40 km.

This problem can be solved in two ways: either to fix one large antenna in one place and move a second smaller antenna (along a railway line, by road or on water), the angles φ being measured as the separation is changed continuously, or to place several moderately-sized antennas in fixed positions and, in this way, obtain several interferometer bases.

It appears that in both cases it would be advantageous to use interferometers with the signals from the two antennas combined by means of a long-distance radio link.

From the above account it is obvious that radioastronomy leads to completely new data on the structure of the universe. It is very important to verify Ryle's results and to extend the observations with large radio telescopes with the aim of obtaining information on radio stars with a lower intensity.

In this work it is important to carry out observations on the size of faint radio stars, as well as to study their spectra (in order to find out whether the liminosity or the density of the sources varies with distance).

In what follows we will discuss several fundamental topics, associated with radio telescope systems, which are of importance in planning investigations of this type.

§ 3. On the "Elongated" Directional Diagram of a Radio Telescope

In order to obtain new information on radio stars it is necessary to work with larger telescopes. As the size of the radio telescopes is increased, it becomes easier to build them with unequal dimensions – with the height smaller than the length along the earth. Let d denote the first length, and D the second, then the directional diagram for such a telescope will not be symmetric but elongated, the ratio of the angular dimensions being D/d.

As was already mentioned, systems with a large ratio D/d were used in carrying out the observations discussed above.

The following question arises in this connection: if the angular dimensions of the directional diagram are $\Delta\varphi_1$ and $\Delta\varphi_2$ ($\Delta\varphi_1 \ll \Delta\varphi_2$), is it possible to increase the resolving power of the equipment also in the direction of the longer axis by moving the pattern lengthwise relatively to the source through a distance of the order of $\Delta\varphi_1$? Let us divide the directional diagram into a number of sections of length $\Delta\varphi$ along the longer axis; we will assume that the directivity factor is constant within the angle $\Delta\varphi$ and that for the various sections it is equal to $F_1, F_2, \ldots, F_n$, where $F_1 < F_2 < \ldots < F_n = 1$.

Let us divide the object being studied into strips of the same width and let $I_1, I_2, \ldots, I_n$ denote, consecutively from the edge, the intensities to be determined. We investigate the one-dimensional case of an extremely elongated diagram. Then, as the pattern is moved a distance Δ relatively to the object, $P_1, P_2, \ldots,$ the energy received, are given by the following expressions (to within a constant factor):

$$P_1 = I_1F_1 \pm \delta_0; \quad P_2 = I_1F_2 + I_2F_1 \pm \delta_0; \qquad (12)$$
$$P_3 = I_1F_3 + I_2F_2 + I_3F_1 \pm \delta_0.$$

It is important to note that we have introduced the quantity δ_0 which determines the error in the measurements due to the limited sensitivity of the apparatus.

From these equations we obtain, step by step,

$$I_1 = \frac{P_1}{F_1} \pm \frac{\delta_0}{F_1}; \qquad I_2 = \frac{P_2}{F_1} - \frac{P_1F_2}{F_1^2} \pm \delta_0 \sqrt{\frac{F_2^2}{F_1^2} + \frac{1}{F_1^2}};$$

$$I_3 = \frac{P_3}{F_1} - \frac{P_2F_2}{F_1^2} + P_1\left[\frac{F_2^2}{F_1^3} - \frac{F_3}{F_1^2}\right] \pm \delta_0 \sqrt{\left(\frac{F_3}{F_1}\right)^2 + \frac{F_2^4}{F_1^6} + \frac{F_2^2}{F_1^4} + \frac{1}{F_1^2}}; \qquad (13)$$

$$I_4 = \frac{P_4}{F_1} - P\left[\frac{F_4}{F_1^2} + 2\frac{F_2F_3}{F_1^3} + \frac{F_2^3}{F_1^4}\right] - P_2\left[\frac{F_2^2}{F_1^2} - \frac{F_2^2}{F_1^3}\right] - P_3\frac{F_2}{F_1^2} \pm$$
$$\pm \delta_0 \sqrt{2\frac{F_2F_3^2}{F_1^6} + \frac{F_3^2}{F_1^4} + \frac{F_4^2}{F_1^4} + \frac{1}{F_1^2} + \frac{F_2^6}{F_1^8} + \frac{F_2^4}{F_1^6} + \frac{F_2^2}{F_1^4}}$$

and so on.

In the above equations we have assumed that the errors combine as the sum of their squares. As can be seen from these expressions, the accuracy of the measurements rapidly decreases as the number of sections, into which the pattern is divided, increases.

Below, we give numerical values for a diagram having the form $(\sin x/x)^2$ when it is divided into various numbers of sections. In the cases when we divide the pattern into 2, 3 and 4 equal sections whose widths are 1/2, 1/3 and 1/4 of the width (at the half-power level) of the directional diagram, we can adopt, with some rounding-off, the following values for the coefficients F_4:

1) $F_1 = 0.15$; $F_2 = 0.8$; $F_3 = 0.6$ two sections,
2) $F_1 = 0.1$; $F_2 = 0.5$; $F_3 = 0.85$ three sections,
3) $F_1 = 0.05$; $F_2 = 0.25$; $F_4 = 0.9$ four sections.

Substituting these values into Equations (13) gives the following expressions for δ:

1) $\delta = 35\,\delta_0$ two sections,
2) $\delta = 250\,\delta_0$ three sections,
3) $\delta = 2500\,\delta_0$ four sections.

As can be seen, even when the diagram is divided into two sections the error of measurement is increased by a factor of 35 and for further division it rapidly increases.

This result shows that there cannot be any useful application of this method in practice because of the enormous loss of sensitivity.

It should be noted that we have investigated a somewhat simplified case, namely, the case when initially the antenna is pointed at empty space and does not receive anything. In fact, during the observation of cosmic radio emission on meter wavelengths there is always a finite initial value for the antenna temperature $T_{\alpha 0}$ produced by the cosmic radio emission. As the antenna is moved, some of the regions of radio emission which were earlier within the directional diagram move outside it, $T_{\alpha 0}$ becomes variable and is no longer constant. This circumstance would make the observations of the type contemplated even more difficult.

§ 4. On the Sensitivity and Resolving Power of a Radio Telescope

In order to develope further the considerations on the diagrams required for large radio telescopes, we have to appraise what we can expect to "see" with a radio telescope of high resolving power. It is of great interest to find approximately the intensities and the number of the sources that will be recorded. We will base our discussion on the data given in paper [1], already mentioned, on the number and the intensities of the discrete sources that are observed at the present time.

For an isotropic spatial distribution of sources, the number of sources which can be observed is given by Equation (2). In the case described by the experimental curve of [1], the exponent $\underline{n}$ has the value -3, instead of $-3/2$, over a fairly wide region.

Therefore, we can write the following general relation which is valid for any values of λ:

$$N_2 = N_1 \left[\frac{I_1 L(\lambda_2)}{I_2 L(\lambda_1)}\right]^n. \tag{14}$$

The smallest measurable intensity for a given telescope is given by the expression:

$$I\min(\lambda) = \frac{2k}{\lambda^2}(N' - 1)\, T_0 \left(\sqrt{\Delta f \tau}\right)^{-1} \Omega_a, \tag{15}$$

where the symbols have their usual meaning.

Comparing two large telescopes, we can estimate N_2, the number of sources that can be recorded with the second telescope, if the number recorded by the first is known:

$$N_2 = N_1 \left(\frac{I_1}{I_2}\right)^n = N_1 \left[\frac{(N_1' - 1)\sqrt{\Delta f_1 \tau_1}\, \Omega_{a1} \lambda_2^2}{(N_2' - 1)\sqrt{\Delta f_2 \tau_2}\, \Omega_{a2} \lambda_1^2}\right]^n \left[\frac{L(\lambda_2)}{L(\lambda_1)}\right]^n. \tag{16}$$

If in both cases the parameters of the receivers are the same and only the area of the antennas is changed, then we find

$$N_2 = N_1 \left(\frac{S_2}{S_1}\right)^n. \tag{17}$$

From the curve of log N as a function of log I given by Ryle, we choose as our initial values the values N_0 and I_0 for which these is still no saturation of the curve. This is the case for

$$\log N_0 = 1.8 \qquad \log I_0 = 0.5,$$

which corresponds to

$$N_0 = 63; \qquad I_0 = 3.2 \cdot 10^{-25}.$$

We note that the total area of the apparatus was about $4 \cdot 10^3$ m^2 and that the weakest sources which were registered had an intensity of $0.7 \cdot 10^{-25}$. If we take I_0 to correspond to $N_0 = 63$, i. e., if we assume that $I_0 = 3.2 \cdot 10^{-25}$, we then have the sensitivity referred to an antenna of area 10^3 m^2.

Using these data, we can now construct a table of the expected number of sources that can be registered by an antenna of another area. We will assume that the sensitivity of the equipment is the same as that of the equipment used to obtain the data used above. The time constant which we use in these calculations remains unchanged and equal to 10-15 sec. To obtain such time constants, the antenna can be fixed if the opening angle of the directional diagram is greater than several minutes of arc.

1	S, м^2	10^3	$4 \cdot 10^3$	$12 \cdot 10^3$	$4 \cdot 10^4$	$8 \cdot 10^4$
2	$\sqrt{S}$, м	33	66	110	200	282
3	$\Delta\varphi$	386	193	127	64	45
4	N_1	63	500	$2.7 \cdot 10^3$	$1.5 \cdot 10^4$	$4.2 \cdot 10^4$
5	N_2	63	$4 \cdot 10^3$	$1.1 \cdot 10^5$	$4 \cdot 10^6$	$3.2 \cdot 10^7$
6	ψ_1	430	190	66	28	17
7	ψ_2	430	54	10	1.7	0.6
8	K_1	1.6	2.1	7.5	10	14
9	K_2	1.6	25	310	$2.8 \cdot 10^3$	10^4
10	$I_{\min}$	$3.2 \cdot 10^{25}$	$0.8 \cdot 10^{-25}$	$0.27 \cdot 10^{-25}$	$0.8 \cdot 10^{-26}$	$0.4 \cdot 10^{-26}$

We proceed to explain the arrangement of the table.

The area of the radio-telescope antenna is denoted by S; the calculations are carried out for a number of cases, starting with $S = 10^3$ to $S = 8 \cdot 10^4$ m^2. The values of $\sqrt{S}$ are given in the second row. These indicate the dimensions (in meters) of a square parabolic reflector required so that its area be S.

The next row 3 gives the resolving power in minutes of arc, $\Delta\varphi$, for a radio telescope of the size indicated in row 2. The wavelength is taken to be 3.7 m in conformity with [1].

Row 4 gives the expected number of sources N_1 per steradian which will have an intensity high enough to be registered by a telescope of area S. Again we base our calculations on Ryle's data. It is assumed that the number of sources with an intensity greater than I increases according to the law with $n = 3/2$ (Equation (17)).

Row 5 gives the expected number of sources N_2 calculated on the assumption that the number increases according to the law with $n = 3$.

Rows 6 and 7 give the average angular separations between the sources (in minutes of arc) for the values of the density given in rows 4 and 5, respectively.

Rows 8 and 9 give the values of the quantities K_1 and K_2 which we will call the "sensitivity nonutilization factors."

We define them as:

$$K_{1,2} = 2\left[\frac{\Delta\varphi}{\psi_{1,2}}\right]^2. \tag{18}$$

Roughly speaking, they indicate how many sources whose intensity is above the threshold $I_{\min}$ fall at the same time into the directional diagram of the radio telescope. The factor 2 is introduced so that when K = 1,

$\Delta\varphi$ would be $\sqrt{2}$ smaller than ψ and it would be possible to resolve most of the sources separated by a distance of $\psi = \sqrt{2}\,\Delta\varphi$.

The minimum intensity of the sources which can be registered is given in Row 10.

Let us discuss some of the results given in this Table.

Turning to the first column of the table, we see that $K_1 = K_2 = 1.6$. The antennas are efficiently used for the finding of the largest number of sources. The departure of K from the value unity can apparently be explained by the fact that Ryle used various combinations of the four antennas and therefore did not have a "saturation" in the curve log N − log I for this region.

However, if n = 3 further on, then even with $S = 4 \cdot 10^3$ m² the nonutilization factor in Ryle's case is about 25. The fact that his curve is highly "saturated" becomes completely understandable. The sensitivity of his telescope is high, while the resolving power is low. Whenever a number of weaker sources fell into the interference pattern, this was interpreted as the observation of a single source.

Further, we see that if n = 3 the factor K_2 becomes extremely large as the area of the telescope is increased.

If n = 3/2, the number of sources that can be registered increases less rapidly; nevertheless the sensitivity of the antenna is not completely utilized.

Moreover, from these figures we see that it is meaningless to propose the construction of large-area telescopes, with tracking of the sources, for the discovery of more sources: the limitations on the observations will be imposed by the resolving power of the apparatus and not by its sensitivity.

Let us now write down the general relations connecting the above quantities.

If the number of sources registered per steradian increases according to the law $N = N_0\,(S/S_0)^n$, then the average angular separation between the sources will be

$$\psi = \frac{1}{\sqrt{N}} = \frac{S_0^{n/2}}{N_0^{1/2}\,S^{n/2}}\,. \tag{19}$$

The opening angle of the antenna diagram is defined as:

$$\Delta\varphi_a = \frac{\lambda}{A}\,; \qquad \Delta\varphi_b = \frac{\lambda}{B}\,; \qquad S = AB. \tag{20}$$

The nonutilization factor is given by the expression:

$$K = 2\frac{\varphi_a\,\varphi_b}{\psi^2} = 2\frac{\lambda^2 N_0}{S_0^n}\,S^{n-1}\,. \tag{21}$$

The values corresponding to n = 3/ 2 and n = 3 are

$$K_1 = 2\frac{\lambda^2 N_0}{S_0^{3/2}}\,S^{1/2}; \qquad K_2 = 2\frac{\lambda^2 N_0}{S_0^2}\,S^2, \tag{22}$$

i. e., the number of sources falling into the directional diagram (irrespective of its size) increases as $S^{1/2}$, or in the worst case, when n = 3, as S^2.

Consequently, in attempting to increase the number of sources registered, there is no point in simply increasing the area of the radio telescope. It is necessary to increase the resolving power by a large factor as the area S is increased.

The calculations carried out above referred to a wavelength of 3.7 m.

Let us now estimate how the results obtained depend on the wavelength λ.

From Equation (16) it follows that

$$N_2(\lambda_2) = N_1(\lambda_1)\left[\frac{(N_1' - 1)\sqrt{\Delta f_2 \tau_2}\,S_2\,L(\lambda_2)}{(N_2' - 1)\sqrt{\Delta f_1 \tau_1}\,S_1 L(\lambda_1)}\right]^n. \tag{23}$$

For a rough estimate we can assume that in the range from decimeter to meter wavelengths there is little variation in the noise level, in view of the fact that at the longer wavelengths the disturbing influence of the general galactic background of radio emission beomes appreciable. Further, we assume that $S_1 = S_2$, $L(\lambda_2)/L(\lambda_1) = (\lambda_2/\lambda_1)^p$, and that the relative width of the receiver transmission-band is $\Delta f/f_0 = \text{const}$.

We then have

$$N(\lambda_2) = N_1(\lambda_1)\left(\frac{\lambda_2}{\lambda_1}\right)^{pn}\left(\sqrt{\frac{\lambda_1 \tau_2}{\lambda_2 \tau_1}}\right)^n. \tag{23a}$$

If the antenna arrangement is fixed, then the time constant is determined by the width of the diagram and, consequently, we can consider that $\tau_1 \equiv \lambda_1$, $\tau_2 \equiv \lambda_2$; in this case, assuming that $\alpha = 0.6$ and $n = 3/2$, we find that the number of radio sources observed is proportional to $\lambda^{0.9}$ but, since the solid angle of acceptance is proportional to λ^{-2}, the nonutilization factor increases with increasing λ as $\lambda^{2.9}$. If we assume that τ is independent of λ, then $N = \lambda^{0.15}$ and the nonutilization factor increases as $\lambda^{2.15}$. For n greater than 3/2, the magnitude of K obviously increases even faster as λ increases. On the other hand, as the wavelength decreases K rapidly falls. Therefore, it is clear that at wavelengths of the order of 1-2 m and shorter, the most important factor for registering the greatest number of sources is the sensitivity, i. e., the area of the antennas.[*]

On the other hand, at wavelengths of the order of 3-10 m the resolving power of the radio telescope becomes the dominant factor.

Let us note, by the way, that since for a given antenna area the number of sources registered will increase as the wavelength increases and since it is easier and cheaper for a telescope of a given area to be built for use at longer wavelengths, it is clear that to accomplish the registration of the greatest number of sources (radio stars) it is desirable to carry out the observations at longer wavelengths.

Let us now investigate what is the most desirable method for carrying out observations at meter wavelengths. We note that the acceptance solid angle of the usual telescope systems (of the type of mirrors, cophase arrays) decreases as S^{-1}. However, there exist other types of antenna systems [3], the "cross-shaped" antennas invented in Australia, which exhibit other types of dependence.

If the smaller pattern-dimension of a symmetric cross-shaped antenna is a and the larger is b, then the solid angle of the pattern is determined by the larger dimension b:

$$(\Delta\varphi)^2 = \frac{\lambda^2}{b^2}; \qquad S = 2ab; \qquad b^2 = \frac{S^2}{4a^2}. \tag{24}$$

If the size of the "cross-shaped" antenna is increased, keeping a constant, then it is obvious that

$$\Delta\varphi^2 = \frac{4\lambda^2 a^2}{S^2}; \quad \Delta\varphi = \frac{2\lambda a}{S}; \tag{25}$$

i. e., the solid angle is inversely proportional to the square and not the first power of the area. Because of this, for a cross-shaped antenna

$$K = N_0 \frac{8\lambda^2 a^2 S^{n-2}}{S_0^n}, \tag{26}$$

and if we take $n = 3/2$, then $K_1 \propto S^{-1/2}$; consequently, for this value of n, the number of sources which fall into the directional diagram will decrease as S increases. This gives us the opportunity of registering them by increasing the area; the sources will not interfere with each other.

If $n = 3$, then $K \propto S$ and, consequently, even in the case of the "cross-shaped" antenna, its area will not be completely utilized.

We note that in all of the discussions it has been assumed that the time constant of the equipment remains constant.

It must be pointed out, however, that in the case when a fixed diagram with a very narrow beam is used, the value of the time constant can be limited as the result of the small transit time of the source across the directional diagram. Then, in order to maintain the former sensitivity, it is necessary to track the source to some extent.

---

[*] If low-noise molecular amplifiers are used, the resolving power of the radio telescope becomes important at even shorter wavelengths.

If $\Delta\varphi$ is the opening angle of the diagram, then the transit time for the source can be found as: $T = \gamma\Delta\varphi$, where γ is a coefficient which depends on the coordinates of the source and the direction of the diagram.

Usually, if $\Delta\varphi$ is expressed in minutes of arc and T in minutes, we can take $\gamma = 0.1$. The time constant τ must be somewhat smaller than T. We can set $\tau \approx 0.3\ T \approx 3\cdot10^{-2}\Delta\varphi$.

For $\Delta\varphi = 10'$, i. e., for an antenna dimension of 300 λ, we obtain a time constant $\tau = 20$ sec, T = 1 min, which is completely adequate for the observations.

For smaller values of $\Delta\varphi$, we must either track the source to some extent or be resigned to some loss of sensitivity.

However (using the average of a number of observations), we can to some extent recover the lost sensitivity in this case also.

The results of the calculations we have made, therefore, show that if we want to utilize intelligently the area of the radio-telescope antenna to discover the largest number of sources, we must use antenna systems of the "cross" type.

We must also make the following remarks in connection with the antenna systems of the "cross"type.

Suppose we have a "cross" with the lengths of the two patterns being a_1 and a_2. Then it is obvious that its directional diagram will be governed by the quantities $\Delta\varphi_1 = \lambda/a_1$, $\Delta\varphi_2 = \lambda/a_2$, and $\Delta\Omega = \Delta\varphi_1\,\Delta\varphi_2 = \lambda^2/a_1a_2$. This is the usual type of "cross-shaped" antenna, only its dimensions in the two directions are different from each other; as the result of this it has an asymmetric directional diagram (Fig. 5, A. Let us now move one of the arms of the cross to a distance b (b > a_2) from the other arm (Fig. 5, B. Then we will obtain a system in the form of an interferometer with a directional diagram whose envelope has the angular dimensions λ/a_1 and λ/a_2; the angular width of one lobe of the diagram is equal to λ/b. n , the number of interference lobes present in the directional diagram, is obviously given by

$$n = \frac{b}{a_1}.$$

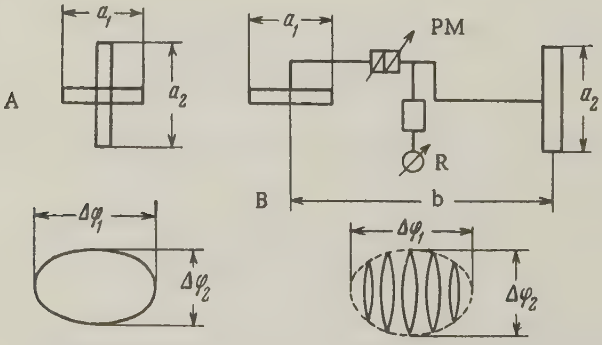


Fig. 5. A) Cross-shaped antenna and its directional diagram; B) spaced-out "cross-shaped" antenna and its directional diagram. PM) Phase modulator; R) receiver. This system is recommended for further investigations of radio stars.

The "spaced-cross" system appears to us to be the most effective for the further study of the distribution of radio stars. It combines, on the one hand, the advantages of a "cross-shaped" antenna by having a very narrow beamwidth and, on the other hand, the advantages of an interferometer which can separate sources with small angular dimensions.

§ 5. On the Magnitude of the Side Lobes in the Directional Diagram of Large Radiotelescopes

In discussing the question of further work on the observation of radio stars, we must evaluate the interference caused by the side lobes in the directional diagram. First of all, let us find out how this interference will affect the observations of the discrete sources of radiofrequency radiation.

Let $\Omega_0, \Omega_1, \Omega_2, \ldots, \Omega_n$ be the solid angles of the zeroth (main) lobe in the diagram, the first side lobe, the second side lobe, etc.

Further, let $F_0, F_1, F_2, \ldots, F_n$ be the corresponding gains for the main lobe and the side lobes. In addition, we assume that for the main lobe $F_0 = 1$.

It must be pointed out that for the usual type of diagram, $(\sin x/x)^2$, the magnitude of Ω_i in a certain region increases approximately as the square of the lobe number, while the values of the F_i are inversely proportional to this number.

The results of measurements made with a given type of telescope can only be reliable if the energy from the weakest source which comes in through the main lobe is greater than the energy from the other sources which comes in through the side lobes, i. e.,

$$P_{\min} \gg F_1 P, \quad P_{\min} \gg F_2 P, \text{ etc.}$$

Let us now find the relative size of those areas of the sky which, because of the nonfulfillment of the above conditions, will not yield reliable information on sources with the lowest intensities.

Let N_0 be the number of sources per steradian which provide a power input $P_{\min}$.

Then N, the number of sources providing a power input P, will be given by:

$$N = N_0 \left(\frac{P_{\min}}{P}\right)^n. \tag{27}$$

Let us assume that $n = 3/2$ for the case of the isotropic distribution and $n = 3$ for Ryle's case. We assume that the interference between one source and another, due to the side lobe of number $\underline{i}$, will occur if $P_{\min} = 10 F_i P$, $P = P_{\min}/10 F_i$; we then find that the number of such sources interfering with the measurements (because the i-th side lobe is present) will be

$$N_i = N_0 (10 F_i)^n. \tag{28}$$

We will now investigate the special cases which are summarized in the tables.

TABLE 1

$n = 3/2$

F_i, %	4.5	1.6	0.83
N_i/N_0	0.3	0.07	0.024
Ω_i/Ω_0	6	10	14
$\frac{\Omega_i/N_i}{\Omega_0 N_0}$	1.8	0.7	0.36

TABLE 2

$n = 3$

F_i, %	4.5	1.6	0.83
N_i/N_0	0.09	$5.1 \cdot 10^{-3}$	$5.7 \cdot 10^{-4}$
$\frac{\Omega_i N_i}{\Omega_0 N_0}$	0.54	0.05	$8 \cdot 10^{-3}$

The results in the first Table have been obtained for the case n = 3/2 ("isotropic" distribution of sources).

The first row contains the dimensions of the side lobes which were used in the calculations. The data refer to the standard diagram $(\sin x/x)^2$ in which there are marked side lobes. The second row contains the values of the ratio N_i/N_0 calculated from Equation (28). The next row gives, for the same diagram, the ratio of the solid angle of the side lobe to that of the main lobe. Finally, the last row contains the values of the product of the quantities in the two preceding rows:

$$M = \frac{\Omega_i N_i}{\Omega_0 N_0}. \tag{29}$$

This ratio is the average number of sources which fall into the diagram (into the i-th side lobe) and which creates an appreciable (≈ 10%) interference with the source whose energy is being measured.

As can be seen, for n = 3/2 the influence of the first lobe is very harmful. In fact it prevents us from carrying out "pure" observations. Similarly, the second lobe and to a smaller extent the third create a considerable amount of trouble.

For the second case when n = 3, this interference is not as strong; nevertheless the first lobe will still spoil about one-half of all the measurements. It is therefore clear that we must decrease the size of the side lobes in order to use radio-telescope antenna systems more efficiently.

This problem must be investigated separately for each antenna system. Some preliminary requirements, however, can be formulated as follows: the height of the first, second, and third lobes must not be greater than 0.5-0.2%, 0.2-0.1%, and 0.1-0.05%, respectively, of the height of the main maximum.

In addition, the values of Ω_i for these lobes must not exceed the values given in the tables. We can obtain side lobes whose size is of this order of magnitude by suitably choosing the distribution of the energy received along the antenna aperture. However, this problem requires careful attention and development.

§ 6. Radio Refractors or Radio Reflectors?

In connection with the problems considered above, we would also like to touch upon the technical problems associated with the creation of large radio telescopes and, in particular, to present some considerations regarding the potentialities of radio reflectors and radio refractors.

In his interesting report, Khaikin [4] divides all radio telescopes, analogously to optical telescopes, into radio reflectors and radio refractors. This division appears to be reasonable. However, we cannot agree with that part of the above paper in which preference is given to radio reflectors and, especially, in which it is stated that "we must consider that the biggest radio telescopes, in the same way as optical telescopes, should be constructed on the principles of reflectors and not refractors."

It is perfectly clear that such an assertion is applicable to optical instruments. Some of the light is invariably lost in passing through the thick lenses and we have no means for amplifying the radiation flux before it becomes attenuated. Because of this, the signal-to-noise ratio (using a term from electronics) decreases.

The situation is different from the above for the radio-frequency region. Here we are able, after reception, first of all to amplify the signals from the separate elements of the refractor and only then send them to the common channel where they are combined. There is no possibility of doing this in the optical case.

It is clear that in dealing with many high-frequency or intermediate-frequency signals (the latter case occurs when converters are placed at each of the refractor elements), we must ensure that the very strict conditions imposed on the amplitudes and, particularly, the phases are fulfilled. This problem is difficult technically, but it can be solved.

In this connection, let us estimate the order of magnitude of the permissible phase-shifts.

For an ordinary interferometer, the transition from the maximum to the minimum of a lobe, i. e., a shift by the width of one lobe, corresponds to a change in the phase difference of 180°.

If we assume that we can tolerate a directional diagram which is stable to within 10%, then, for the case of two antennas, the permissible phase-shift is 18°. If the number of separate elements in the radio refractor is increased, then from simple considerations it is clear that if the phase-shift produced by each element does not exceed 18° then there will be no appreciable deviation of the diagram (within 10%) from its initial position. But this phase stability can be achieved in practice, even if special measures have to be taken. As regards the conditions imposed on the amplitudes, a spread of 10% is also permissible. This accuracy can be achieved in practice.

Thus, even if the main difficulty in creating radio refractors, the correct combination of the signals from the separate elements, presents great technical problems, it can already be overcome at the present time. Today, we can plan and construct telescopes of the refractor type several kilometers in size with a symmetrical directional diagram which is electrically controlled. On the other hand, it appears to be completely impossible to build radio reflectors of this type which would have a similar directional diagram. The "Mills Cross" is the most remarkable radio-refractor system and, at the present time, there does not appear to be any possibility for replacing it by an equivalent reflector system.

The following important advantage of radio refractors as compared to radio reflectors must be noted. In the case of radio refractors, where we have amplifying (or mixing) elements placed in the circuit after the separate parts of the antenna system, we are able to obtain various reception characteristics for the radio telescopes by connecting these elements in various combinations.

For example, in the case of a "cross-shaped" antenna consisting of a number of elements, we can combine groups of elements in various ways to obtain interferometer arrays of various types, interferometers with various base lines, etc.

In connection with the above account, it seems to us that the future of the largest radio telescopes belongs to radio refractors.

P. N. Lebedev Physical Institute
USSR Academy of Sciences

Received October 6, 1956

LITERATURE CITED

[1] M. Ryle, Observatory 75, 887, 137.

[2] M. Ryle and P. A. G. Scheuer, Proc. Roy. Soc. 1183, 230 (1955).

[3] B. Mills, Observatory (October, 1953).

[4] S. E. Khaikin, Proc. of the 5th Conference on Cosmogony [in Russian] (Izd. AN SSSR, 1956), p. 9.

ELECTROMAGNETIC RADIATION BY COSMIC-RAY PARTICLES IN THE GALAXY

A. A. Korchak

The spectrum of the electromagnetic emission by cosmic ray particles in the galactic magnetic field has been calculated, assuming that the latter is nonuniform and that the character of the particle trajectory is arbitrary. It is shown that the spectrum coincides with the observed spectrum of the spherical component of radio emission if some assumptions are made about the lower end of the energy spectrum of the particles. A more precise estimate is made of the intensity of the emission by the electron component.

At the present time, it has been established that cosmic radio emission consists of two components: a short-wave component of centimeter wavelengths and a long-wave component of wavelength longer than a meter [1, 2]. The radiation on the short wavelengths is strongly concentrated towards the galactic plane and its intensity is independent of the frequency, which enables it to be identified with the emission from the ionized interstellar gas. The long-wave component is localized in a region with dimensions of the order of 15 kps and has a falling frequency spectrum, $I - v^{-\alpha}$, where $0.5 \leq \alpha \leq 1$ [3]. Sometimes on meter wavelengths the isotropic extragalactic component is distinguished [4, 5].

The spherical component of the radio emission is most readily explained by the radiation emitted by the electron component of the cosmic rays in the galactic magnetic fields. This hypothesis was proposed in 1950 by Alfven and Herlofson [6] and Kiepenheuer [7] and was then developed by Getmantsev and Ginsburg [8-12].

For it to be successful, it was first of all necessary to show that the spectrum of the synchrotron radiation of the primary cosmic-ray component agreed with the observed spectrum of the radiofrequency radiation. This was shown to be the case, simultaneously, in references [10] and [13]. However, in these calculations the basic assumptions were that, in galactic space, the cosmic-ray particles move along circular trajectories in a uniform magnetic field, which of course is an approximation. It will be shown below that this assumption is not an essential one. With a reasonable assumption as to the lower end of the energy spectrum of the cosmic rays, their radiation spectrum does not depend on the distribution of the magnetic-field intensity, H, in the quasi-uniform regions of the galactic magnetic field, nor on the distribution of the angles between the direction of particle motion and the direction of the magnetic field.

The Radiation Spectrum of the Cosmic Rays.

Let us initially consider the simplest case when all of the cosmic rays move along circular trajectories in a uniform magnetic field H_c*. It is known that in this case electromagnetic emission will take place; the power radiated by an ultra-relativistic particle, $p(\nu)$, in a unit frequency range at the frequency ν is given by the expression [14]:

---

*In the following we will take H_c equal to the average field of the Galaxy.

$$p(\nu E) = \frac{4\pi}{\sqrt{3}} \frac{z^2 e^2 m^2 c^3 \nu}{E^2} \int\limits_{\nu/\nu_c}^{\infty} K_{5/3}(x)\,dx, \tag{1}$$

where m, ze, E are the mass, charge and energy of the particle, $\nu_c = \frac{3zeH_c}{4\pi mc}\left(\frac{E}{mc^2}\right)^2$ is the frequency at which the maximum energy is radiated and $K_{5/3}(x)$ is the Bessel function of an imaginary argument. If we use the observed spectrum of the primary cosmic rays

$$dN(E) = kE^{-\gamma}dE, \tag{2}$$

where $k = (\gamma - 1) N_0 E_0^{\gamma-1}$ is a normalization constant, N_0 is the average number of cosmic-ray particles in a unit volume of galactic space and E_0 is some minimum energy in the spectrum, then the total radiation of a unit volume of galactic space in the unit frequency range at ν is

$$P(\nu) = \int\limits_{E_0}^{\infty} p(\nu E)\,dN(E) = \frac{4\pi}{\sqrt{3}} z^2 e^2 m^2 c^3 k \nu \int\limits_{E_0}^{\infty} E^{-(\gamma+2)} dE \int\limits_{\nu/\nu_c}^{\infty} K_{5/3}(x)\,dx. \tag{3}$$

Let us introduce a new variable $y = \nu/\nu_c$; then

$$P(\nu) = \frac{\sqrt{3}}{2} k \frac{z^3 e^3 H_c}{mc^2} \left(\frac{3zeH_c}{4\pi m^3 c^5}\right)^{(\gamma-1)/2} I(\nu)\, \nu^{-(\gamma-1)/2}, \tag{4}$$

where

$$I(\nu) = \int\limits_0^{y_0} y^{(\gamma-1)/2} dy \int\limits_{y_0}^{\infty} K_{5/3}(x)\,dx,$$

$$y_0 = \frac{\nu}{\nu_c^0}, \quad \nu_c^0 = \frac{3zeH_c}{4\pi mc}\left(\frac{E_0}{mc^2}\right)^3. \tag{5}$$

It is obvious that

$$I(\nu) = \int\limits_0^{\infty} y^{(\gamma-1)/2} dy \int\limits_y^{\infty} K_{5/3}(x)\,dx - \int\limits_{y_0}^{\infty} y^{(\gamma-1)/2} dy \int\limits_y^{\infty} K_{5/3}(x)\,dx = I_1 - I_2. \tag{6}$$

The first integral on the right-hand side does not depend on the frequency and can be easily calculated:

$$I_1 = \int\limits_0^{\infty} y^{(\gamma-1)/2} dy \int\limits_y^{\infty} K_{5/3}(x)\,dx = \int\limits_0^{\infty} K_{5/3}(x)\,dx \int\limits_0^{x} y^{(\gamma-1)/2} dy =$$

$$= \frac{2}{\gamma+1} \int\limits_0^{\infty} x^{(\gamma+1)/2} K_{5/3}(x)\,dx = \frac{2^{(\gamma+1)/2}\Gamma\left(\frac{3\gamma-1}{12}\right)\Gamma\left(\frac{3\gamma+19}{12}\right)}{\gamma+1}, \tag{7}$$

where Γ is the Euler gamma function. The second integral, in the range of frequencies $10^7 - 10^9$ cps (30 cm – 30 m) which are of interest to us, can be neglected if a certain condition is imposed on E_0. In fact, let E_0 be such that $y_0 = \frac{\nu}{\nu_d^0} = \frac{4\pi mc\nu}{3zeH_c}\left(\frac{mc^2}{E_0}\right)^2 \gg 1$; then in Equation (6) because $x \geq y \geq y_0 \gg 1$ we can put $K_{5/3}(x) = \sqrt{\frac{\pi}{2x}}\, e^{-x}[1 + O(1/x)] \sim \sqrt{\frac{\pi}{2x}}\, e^{-x}$, which is valid for large values of the argument x. Therefore

$$\int_{y}^{\infty} K_{5/3}(x)\,dx \sim \sqrt{\frac{\pi}{2}}\, y^{-1/2} e^{-y}$$

and

$$I_2 = \int_{y_0}^{\infty} y^{(\gamma-1)/2} dy \int_{y}^{\infty} K_{5/3}(x)\,dx \sim \sqrt{\frac{\pi}{2}}\, y_0^{(\gamma-2)/2} e^{-y_0} \ll 1, \tag{8}$$

while at the same time I_1 with $\gamma = 1, 2, 3$ takes the values 5, 2.1, 1.6. Thus, with the assumption made that $y_0 \gg 1$, $I \sim I_1$ is independent of the frequency and

$$P(\nu) = A\nu^{-(\gamma-1)/2}, \tag{9}$$

where*

$$A = \frac{\sqrt{3}}{\gamma+1}\Gamma\left(\frac{3\gamma-1}{12}\right)\Gamma\left(\frac{3\gamma+19}{12}\right) k\, \frac{z^3 e^3 H_c}{mc^2}\left(\frac{3zeH_c}{2\pi m^3 c^5}\right)^{(\gamma-1)/2} \tag{10}$$

Let us now investigate the more general case of a nonuniform field. For an arbitrary motion of an ultra-relativistic particle in an electromagnetic field, instead of (1) we have [15]

$$p(\nu ER) = \sqrt{3}\,\frac{z^2e^2}{R}\left(\frac{E}{mc^2}\right)\frac{\nu}{\nu_c}\int_{\nu/\nu_c}^{\infty} K_{5/3}(x)\,dx, \tag{11}$$

where R is the instantaneous radius of curvature of the trajectory, $\nu_c = \frac{3}{4\pi}\frac{c}{R}\left(\frac{E}{mc^2}\right)^3$ is the critical frequency and all the other symbols are as before. The equation is correct for $\nu \gg \frac{e}{2\pi R}$, which always holds in the case that we are considering. For the ultra-relativistic case and in the presence of a magnetic field only

$$R = \frac{cp_t}{zeH} = \frac{E}{zeH\sin\theta} \text{ and } \nu_c = \frac{3zeH\sin\theta}{4\pi mc}\left(\frac{E}{mc^2}\right)^2, \tag{12}$$

where θ is the angle between the particle velocity and the direction of the magnetic field and p_t is the projection of the momentum in a plane perpendicular to H.

Let us introduce the function $W_2(H)$, the distribution function of the average magnetic-field intensity in the quasi-uniform regions of the galactic magnetic field, the distribution function $\theta W_3(\theta)$ of the angles of the cosmic-ray particles and the distribution function $W_1(E)$ of their energies. If we assume that these distributions are independent, then the spectral power radiated by a unit volume of the Galaxy is

$$P(\nu)\int_0^{\pi} W_3(\theta)\,d\theta \int_0^{H_m} W_2(H)\,dH \int_{E_0}^{\infty} W_1(E)\,p(\nu E\theta H)\,dE =$$

---

*Equations (9) and (10), given earlier in Reference [13], are the same as the corresponding equations in References [11] and [12] where the power radiated by the primary cosmic rays is expressed by means of a certain function which was introduced and tabulated by Vladimirskii [18].

The difference between the numerical coefficients, arising from the approximate nature of the tables, is not significant. As can be seen from the text it is not necessary to introduce a special function when calculating the radiation spectrum.

$$= \frac{4\pi}{\sqrt{3}} z^2e^2m^2c^3\nu \int_0^{\pi} W_3(\theta)\,d\theta \int_0^{H_m} W_2(H)\,dH \int_{E_0}^{\infty} W_1(E)E^{-2}dE \int_{\nu/\nu_c(E\theta H)}^{\infty} K_{5/3}(x)\,dx. \tag{13}$$

Substituting the experimentally-known function $W_1(E) = kE^{-\gamma}$ (see (2)) and introducing the new variable $y = \nu/\nu_c$, we obtain

$$P(\nu) = \frac{\sqrt{3}}{2^{(\gamma+1)/2}} \frac{kz^3e^3}{mc^2} \left(\frac{3ze}{2\pi m^3c^5}\right)^{(\gamma-1)/2} \nu^{-(\gamma-1)/2} \int_0^{\pi} W_3(\theta) \sin^{(\gamma+1)/2}\theta\, d\theta \times$$

$$\times \int_0^{H_m} W_2(H)\, H^{(\gamma+1)/2} dH \int_0^{y_0'} y^{(\gamma-1)/2} dy \int_y^{\infty} K_{5/3}(x)\,dx, \tag{14}$$

where

$$y_0' = \frac{\nu}{\nu_c^{0'}}, \quad \nu_c^{0'} = \frac{3zeH\sin\theta}{4\pi mc}\left(\frac{E}{mc^2}\right)^2.$$

Since $0 \leq \sin\theta \leq 1$ and $0 \leq H \leq H_m \sim H_c$, then

$$y_0' > y_0 = \frac{4\pi m^3c^5\nu}{3zeH_cE_0^2}.$$

Hence, with the same assumption that $y_0 \gg 1$,

$$\int_0^{y_0'} y^{(\gamma-1)/2} dy \int_y^{\infty} K_{5/3}(x)\,dx \approx I_1 = \frac{2^{(\gamma+1)/2}\Gamma\left(\frac{3\gamma-1}{12}\right)\Gamma\left(\frac{3\gamma+19}{12}\right)}{\gamma+1} \tag{15}$$

and

$$P(\nu) = A'\nu^{-(\gamma-1)/2} \tag{16}$$

where

$$A' = \frac{\sqrt{3}}{\gamma+1}\Gamma\left(\frac{3\gamma-1}{12}\right)\Gamma\left(\frac{3\gamma-19}{12}\right) I_2I_3k\frac{z^3e^3}{mc^2}\left(\frac{3ze}{2\pi m^3c^5}\right)^{(\gamma-1)/2},$$

$$I_2 = \int_0^{H_m} W_2(H)\, H^{(\gamma+1)/2} dH, \quad I_3 = \int_0^{\pi} W_3(\theta) \sin^{(\gamma+1)/2}\theta\, d\theta. \tag{17}$$

Therefore, with the same choice of E_0 the spectrum of the radiation of the cosmic rays does not depend on the distribution function $W_2(H)$, nor on the distribution of the angles θ.

Let us compare (9) and (16). It is easily seen that for a uniform field, H_c, and circular particle trajectories

$$W_3(\theta) = \delta(\theta - \pi/2),\ W_2(H) = \delta(H - H_c),\ I_2 = H_c^{(\gamma+1)/2},\ I_1 = 1,\ A' = A$$

and (16) becomes the same as (9). In all other cases $I_2 \sim H_c^{(\gamma+1)/2}$, $I_3 < 1$ and $A' < A$. In particular, for a uniform distribution of the cosmic-ray particles with respect to the angles θ, $W_3(\theta) = 1/2$ and the averaging with respect to the angles θ gives*

---

* The possibility of the appearance of a factor 2/3 due to averaging was pointed out in Ref. [9], p. 356.

$$I_3 = \frac{1}{2}\int_0^{\pi} \sin^{(\gamma+3)/2}\theta\, d\theta = 2^{(\gamma+1)/2}\, \frac{\Gamma^2\left(\frac{\gamma+5}{4}\right)}{\Gamma\left(\frac{\gamma+5}{2}\right)} \approx 0.7 \tag{18}$$

with $\gamma = 2-3$ and $A' \approx {}^2/_3\, A$.

Comparison with Observations

For a comparison with observation we equate the radiation emitted by the electron component of the cosmic rays $I = \frac{1}{4\pi} P(\nu)\, l = \frac{A' l}{4\pi}\, \nu^{-(\gamma-1)/2}$ to the observed intensity of the cosmic radio emission $I = a\nu^{-\alpha}$, where $0.5 \leq \alpha \leq 1$ and $a = 2 \cdot 10^{-13}\ \frac{\text{erg}}{\text{cm}^2 \cdot \text{sec} \cdot \text{cps}^{1/2} \cdot \text{sterad}}$ when $\alpha = 1/2$ and $a = 2 \cdot 10^{-9}\ \frac{\text{erg}}{\text{cm}^2\, \text{sec}\, \text{sterad}}$ when $\alpha = 1$ according to the data of Reference [16]. From this it follows that the electron component must have a power spectrum with $\gamma = 2\alpha + 1 = 2 - 3$, which is close to the observed spectrum of the proton-nuclear component.* The supplementary condition $y_0 = \nu/\nu_c^0 \gg 1$, taken by us in the calculation of the spectrum, in the present case will mean that the cosmic-ray spectrum must extend to low energies at least down to $E_0 = 5 \cdot 10^7$ ev (Table 1).

TABLE 1

Electron energy E, ev	10^6	10^7	10^8	10^9	10^{10}
ν_c in cps with $H = 10^{-5}$ gauss	$2 \cdot 10^2$	$2 \cdot 10^4$	$2 \cdot 10^6$	$2 \cdot 10^8$	$2 \cdot 10^{10}$
ν_c in cps with $H = 10^{-4}$ gauss	$2 \cdot 10^3$	$2 \cdot 10^5$	$2 \cdot 10^7$	$2 \cdot 10^9$	$2 \cdot 10^{11}$

From the comparison it also follows that $A' l/4\pi = a$, which enables the evaluation of k using (17) and, consequently, the evaluation of the density of electrons with energies greater than $5 \cdot 10^7$ ev necessary to explain the radio emission. Inasmuch as observations give [17] the maximum possible abundance of the electron component for an energy greater than 10^9 ev, to make a comparison with observation it is convenient to give a table of the density of electrons with energies higher than 10^9 ev required for various assumptions in regard to H_c and l.. The data given in Table 2 do not differ from those given in References [9-12] to within an order of magnitude.**

TABLE 2

	$\alpha = 0{,}5,\ \gamma = 2$			$\alpha = 1,\ \gamma = 3$		
	$H = 5 \cdot 10^{-5}$ gauss	$H = 5 \cdot 10^{-6}$ gauss	$H = 10^{-6}$ gauss	$H = 10^{-5}$ gauss	$H = 5 \cdot 10^{-6}$ gauss	$H = 10^{-6}$ gauss
l—10^{22} cm	$1.2 \cdot 10^{-11}$	$3.3 \cdot 10^{-11}$	$3.8 \cdot 10^{-10}$	$5.5 \cdot 10^{-12}$	$2.2 \cdot 10^{-11}$	$5.5 \cdot 10^{-10}$
l—$5 \cdot 10^{22}$ cm	$2.3 \cdot 10^{-12}$	$6.6 \cdot 10^{-12}$	$7 \cdot 10^{-11}$	$1.1 \cdot 10^{-12}$	$4.4 \cdot 10^{-12}$	$1.1 \cdot 10^{-10}$
l—10^{23} cm	$1.2 \cdot 10^{-12}$	$3.3 \cdot 10^{-12}$	$3.8 \cdot 10^{-11}$	$5.5 \cdot 10^{-13}$	$2.2 \cdot 10^{-12}$	$5.5 \cdot 10^{-11}$

---

* See also Reference [12] where the reason for the possible difference of the spectra is also investigated.

** In References [9] and [12] the intensity per min was used for the comparison; therefore with $\alpha = 1$, a was taken as $a = 1.5 \cdot 10^{-10}\ \frac{\text{erg}}{\text{cm}^2\, \text{sec}\, \text{sterad}}$. From this the condition $3.1 \cdot 10^{-15}\ KRH^2 \approx 5 \cdot 10^{-21}$ was obtained, which, with $H \sim 10^{-5}$ oersted and $R \sim 5 \cdot 10^{22}$ cm, gives $k \sim 10^5\ \frac{(\text{ev})^2}{\text{cm}^3}$. Then, as a safety factor, k was increased by an order of magnitude, i. e., to $10^6\ (\text{ev})^2/\text{cm}^3$ (see Reference [9], Equations (13) and (16)). If it is not assumed that $H > 10^{-5}$ and $R > 5 \cdot 10^{22}$ cm, then this means that a was taken as an order of magnitude greater, i. e., $1.5 \cdot 10^{-9}$ erg/(cm^3 sec sterad). On the basis of [16] we have taken a somewhat higher value above for the intensity at the maximum.

The slight discrepancy by a factor of approximately two is associated with the appearance of the factor 2/3 as the result of the averaging and the use of a somewhat larger value of a, since the comparison was made with radio-emission intensity at the maximum. The conclusion drawn in References [9-12], that the spherical component of the radio emission can be explained by the radiation emitted by the electron component of the cosmic rays, remains unaltered; however, a more accurate determination of the upper limit of its density will be an important check of the hypothesis.

M. V. Lomonosov Moscow State University

Received November 3, 1956

LITERATURE CITED

[1] I. S. Shklovskii, Astron. J. (USSR) 29, 418 (1952).

[2] I. S. Shklovskii, Astron. J. (USSR) 30, 15 (1953).

[3] B. Lovell and J. Clegg, Radio Astronomy [Russian translation] (ITTL, Moscow, 1953).

[4] J. L. Pawsey and R. N. Bracewell, Radio Astronomy (Oxford, 1955).

[5] I. S. Shklovskii, Astron. J. (USSR) 30, No. 5 (1953).

[6] H. Alfven and N. Herlofson, Phys. Rev. 78, 616 (1950).

[7] K. O. Kiepenheuer, Phys. Rev. 79, 738 (1950).

[8] V. L. Ginzburg, Doklady Akad. Nauk SSSR 76, 377 (1951).

[9] V. L. Ginzburg, Uspekhi Fiz. Nauk 51, 343 (1953).

[10] G. G. Getmantsev, Doklady Akad. Nauk SSSR 83, 557 (1952).

[11] V. L. Ginzburg and M. I. Fradkin, Doklady Akad. Nauk SSSR 42, No. 3 (1953).

[12] V. L. Ginzburg, G. G. Getmantsev and M. I. Fradkin, Proceedings of the 3rd Conference on Cosmogony, [in Russian] (Izd. AN SSSR, 1954).

[13] A. A. Korchak and Ia. P. Terletskii, J. Exptl.-Theoret. Phys. (USSR) 22, 507 (1952).

[14] L. D. Landau and E. M. Lifshits, Theory of Fields [in Russian] (Gostekhizdat, Moscow-Leningrad, 1948) Section 73; D. Ivanenko and A. Sokolov, Classical Theory of Fields [in Russian] (GITTL, Moscow-Leningrad, 1951) Section 43.

[15] J. Schwinger, Phys. Rev. 75, 1912 (1949).

[16] J. Piddington, Monthly Notices Roy. Australian Soc. 3, 45 (1951).

[17] C. L. Critchfield, E. P. Ney and S. Oleksa, Phys. Rev. 85, 461 (1952).

[18] V. V. Vladimirskii, J. Exptl.-Theoret. Phys. (USSR) 18, 392 (1948).

THE THEORY OF THE FLUCTUATIONS OF THE OBSERVED NUMBER OF GALAXIES

T. A. Agekian

Up to the present either the influence of the patchy structure of absorbing matter or that of the tendency of galaxies to form clusters have been considered when explaining the observed fluctuations in the number of galaxies per unit area of the sky. In reality we have to deal with the combined action of both of these factors. In the present paper V. A. Ambartsumian's equation is generalized by assuming that radiating and absorbing matter, on the one hand, are each distributed in space according to some law without fluctuations; on the other hand, they form patches, also distributed according to some law but with inherent fluctuations. The equations are solved by the method of moments and applied to fluctuations in the number of galaxies. The solution, which accounts for the action of these two factors, shows that the parameters characterizing the structure of dark matter can be separated from the parameters determining the distribution of galaxies in the Metagalaxy. The theory is applied to the counts of galaxies made by Hubble, Shapley and Shane and Wirtanen. The result shows that the predominant number of galaxies in the Metagalaxy belong to small clusters, with the mean number of nondwarf galaxies $\overline{g} = 8.3 \pm 1.6$. The result $\overline{g} < 287$ derived by Neyman, Scott and Shane, using the counts made by Shane and Wirtanen, is incorrect and is explained by the fact that the patchy structure of absorbing matter was not taken into account. Small groups of galaxies like the local system of galaxies, and not the gigantic clusters, as the cluster in Virgo, Corona Borealis, etc., are characteristic of the Metagalaxy. The following were also derived: the mean absorption in one cloud, which for the above-mentioned counts is: $0^m.46$, $0^m.39$, $0^m.24$, and the mean absorption in the direction of the galactic pole: $0^m.39$, $0^m.26$, $0^m.85$.

1. Derivation of the Theoretical Relations

The number of galaxies per unit area of the sky, as is known, is a number which is subject to fluctuations which considerably exceed the inherent changes. To explain this phenomenon two hypotheses have been put forward: 1) the observed fluctuations are due to the nature of the distribution of light-absorbing matter in the galaxy; 2) the observed fluctuations are due to the tendency of galaxies to form clusters.

The first hypothesis was investigated by V. A. Ambartsumian in References [1, 2], which gave rise to the theory of fluctuations in stellar statistics. The second hypothesis was developed by Neyman and Scott [3, 4] and by other authors.

In both of these treatments the observed fluctuations were explained by one of these mechanisms, while the possible influence of the other was not taken into account. At the present time, however, it can be considered established in fact both that the light-absorbing matter in the galaxy has a nonuniform structure and that the majority of the galaxies are members of clusters of galaxies. It is therefore clear that the observed fluctuations in the distribution of galaxies are produced by both causes. It is necessary to investigate simultaneously the influence of these effects and, if possible, to separate them.

In Reference [5] we have derived a generalization of V. A. Ambartsumian's equation for the case of a varying field. For the solution of this problem we will write these equations in an even more general form. We will consider that both the dark matter and the light-emitting matter are not only collected into separate clouds and distributed according to some law with inherent fluctuations, but that they are also distributed in space continually according to some law without fluctuations.

Let $\underline{x}$ be the light intensity observed at a given point, in a given direction, $\psi_s(x)$ be the probability to observe from the point $\underline{s}$, in a given direction, an intensity not exceeding $\underline{x}$, τ the optical thickness of the dark matter lying along a unit length of path, η the radiation from a unit volume of light-emitting matter in a unit solid angle, κ the probability of encountering a dark cloud in a unit path length, $\underline{q}$ the coefficient of transmission of the cloud, $\varphi(q)$ the normalized distribution function of $\underline{q}$, λ the probability for encountering a concentration of the light-emitting matter in a unit path length, ξ the radiation emitted into a unit solid angle by the concentration of matter lying in a cylinder whose axis coincides with the line of sight and whose base is of unit area and $\chi(\xi)$ the normalized distribution function of ξ.

We will consider that the functions $\varphi(q)$ and $\chi(\xi)$ are independent of the position of the point; all of the other quantities depend on the position of the point.

Using the generalized invariance principle for the transition from the point $s + \Delta s$ to the point $\underline{s}$ on the line of sight, in the same way as it was done in Reference [5], we derive the following equation:

$$\psi_{s+\Delta s}(x + x\tau\Delta s - \eta\Delta s)(1 - \varkappa\Delta s - \lambda\Delta s) + \varkappa\Delta s \int\limits_{(q)} \psi_{s+\Delta s}\left(\frac{x}{q}\right)\varphi(q)\,dq +$$
$$+ \lambda\Delta s \int\limits_{(\xi)} \psi_{s+\Delta s}(x-\xi)\,\chi(\xi)\,d\xi = \psi_s(x). \tag{1}$$

Keeping in the first term only the quantities of the first order in Δs and going to the limit $\Delta s \to 0$, we obtain

$$\frac{d\psi_s(x)}{ds} - (\varkappa + \lambda)\psi_s(x) + (x\tau - \eta)\psi_s'(x) +$$
$$+ \varkappa \int\limits_{(q)} \psi_s\left(\frac{x}{q}\right)\varphi(q)\,dq + \lambda \int\limits_{(\xi)} \psi_s(x-\xi)\,\chi(\xi)\,d\xi = 0. \tag{2}$$

If by differentiation we go over to the differential distribution function, f(x), then multiplying the equation obtained by $x^n dx$ and integrating from 0 to ∞, in view of

$$\int\limits_0^\infty [xf(x)]' x^n dx = x^{n+1} f(x)\Big|_0^\infty - n\int\limits_0^\infty x^n f(x)\,dx = -n\mu_n,$$

$$\int\limits_0^\infty x^n dx \int\limits_{(\xi)} f(x-\xi)\,\chi(\xi)\,d\xi = \int\limits_{(\xi)} \chi(\xi)\,d\xi \int\limits_{-\xi}^\infty (x+\xi)^n f(x)\,dx =$$

$$= \int\limits_{(\xi)} \chi(\xi)\,d\xi \int\limits_0^\infty \sum_{k=0}^n C_n^k x^{n-k}\xi^k f(x)\,dx = \sum_{k=0}^n C_n^k \bar{\xi}^k \mu_{n-k} = \mu_n + \sum_{k=1}^n C_n^k \bar{\xi}^k \mu_{n-k},$$

we find

$$\frac{d\mu_n(s)}{ds} - [\varkappa(1-\overline{q^n}) + n\tau]\,\mu_n(s) + n\eta\mu_{n-1}(s) + \lambda\sum_{k=1}^{n} C_n\overline{\xi^k}\mu_{n-k}(s) = 0, \tag{3}$$

where $\mu_i(s)$ are the i-th order moments of the function f(x) at the point s.

Solving the differential Equation (3) for $\mu_n(s)$, the recurrence formula for the central moments can be obtained.

We will consider particular cases which have a practicsl value.

In investigating the fluctuations of the brightness of the Milky Way we can consider that the stars are distributed according to a definite law without fluctuations, while the observed fluctuations in the brightness are produced only by the chance distribution of dark clouds. The dark substratum is absent. Then $\lambda \equiv 0$, $\tau \equiv 0$ and we obtain the following equation:

$$\frac{d\mu_n(s)}{dt} - \varkappa(1-\overline{q^n})\mu_n(s) + n\,\eta\mu_{n-1}(s) = 0,$$

which has been investigated in Reference [5].

Let us now consider the fluctuations in the brightness of the Metagalaxy which would be seen by an observer situated at the edge of our galaxy. As is known the galaxies form clusters and multiple galaxies and are also found as separate galaxies. These formations we will consider to be concentrations of light-emitting matter. Therefore, $\lambda \neq 0$ and $\eta \equiv 0$. Very little is known about the dark matter in the Metagalaxy. Let us assume that it is distributed continuously according to some law without flucturations, while clouds of it are absent, i. e., $\tau \neq 0$, $\kappa \equiv 0$. Equation (3) will then have the form:

$$\frac{d\overset{*}{\mu}_n(s)}{ds} - n\tau(s)\,\overset{*}{\mu}_n(s) + \lambda(s)\sum_{k=1}^{n} C_n^k\overline{\xi^k}\overset{*}{\mu}_{-k}(s) = 0. \tag{4}$$

To avoid Olbert's paradox, which in our case will lead to the moments of the function f(x) becoming infinite, it is necessary to assume either, that as $s \rightarrow \infty$, λ tends to zero faster than $1/s^{1+\sigma}$, where σ is a fixed, arbitrarily small positive quantity, or that in this case τ does not tend to zero too rapidly.

In view of the fact that, in what follows, we intend to go over from the fluctuations in the brightness of the Metagalaxy to the fluctuations in the number of galaxies down to a certain magnitude, we can assume that λ is equal to zero for all $s > R$, since all the galaxies down to a certain magnitude will be contained in a sphere of a certain radius R. This will guarantee the elimination of Olbert's parodox. In view of the small role played by the metagalactic absorption of light, for not too large distances τ can be put equal to zero.

In this case the equation

$$\frac{d\overset{*}{\mu}_n(s)}{ds} + \lambda(s)\sum_{k=1}^{n} C_n^k\overline{\xi^k}\overset{*}{\mu}_{n-k}(s) = 0, \tag{5}$$

because $\overset{*}{\mu}_n(s) \rightarrow 0$ as $s \rightarrow R$, has the recurrent solution

$$\overset{*}{\mu}_n(s) = \sum_{k=1}^{n} C_n^k\overline{\xi^k}\int_{s}^{R} \lambda\overset{*}{\mu}_{n-k}(s)ds. \tag{6}$$

If we now assume that all the galaxies have the same absolute magnitude, while the mean density of their clusters inside a sphere of radius R is everywhere constant, then for all $s \leq R$ we have λ = const and for s = 0 Equation (6) will have the form:

$$\left.\begin{array}{l}\mu_1^*(0) = \lambda R\overline{\xi};\\ \mu_2^*(0) = \lambda^2 R^2\,\overline{\xi}^2 + \lambda R\overline{\xi^2};\\ \mu_3^*(0) = \lambda^3 R^3\overline{\xi}^3 + 3\lambda R^2\,\overline{\xi\xi^2} + \lambda R\overline{\xi^3}.\\ \cdots\cdots\cdots\cdots\cdots\cdots\end{array}\right\}, \tag{7}$$

which is the solution of the particular problem considered.

Let us now investigate the fluctuations in the brightness of the Metagalaxy which are observed at points situated inside the Galaxy. Since, we are considering the brightness of the Metagalaxy (in the following we consider the case of the number of observed galaxies) and not of the Galaxy, then in the passage of a light ray inside the Galaxy $\eta \equiv 0$, $\lambda \equiv 0$. We will consider that all the dark matter in the Galaxy is collected into clouds. Therefore, $\kappa \neq 0$, $\tau \equiv 0$.

Equation (3) will take the form:

$$\frac{d\mu_n(s)}{ds} - \varkappa(s)\left(1 - \overline{q^n}\right)\mu_n(s) = 0. \tag{8}$$

In solving this equation, it is necessary to keep in mind the fact that in going outside the Galaxy as s increases, we arrive at the problem, already investigated, of the fluctuations in the brightness of the Metagalaxy observed from the edge of our Galaxy. Therefore, taking into account the difference in the scale of the galaxy and Metagalaxy, we can write:

$$\mu_n(\infty) = \mu_n^*(0).$$

Consequently, solving (8), we obtain

$$\mu_n(s) = \mu_n^*(0)\, e^{-(1-\overline{q^n})\int\limits_s^\infty \varkappa(s)ds}. \tag{9}$$

If we assume that κ (s) is a function only of the distance from the plane of the Galaxy, then

$$\int\limits_0^\infty \varkappa(s)\,ds = \zeta \operatorname{cosec} b,$$

where b is the galactic latitude and ζ is the mathematical expectation of the number of dark clouds in the direction of the galactic pole. From Equation (9) it then follows, with s = 0,

$$\mu_n(0) = \mu_n^*(0)\, e^{-(1-\overline{q^n})\zeta \operatorname{cosec} b}. \tag{10}$$

In order to go over to the case of the observed number of galaxies, N, it is necessary to take into account that because of screening by the dark matter

$$N \sim x^{3/2},$$

while because of fluctuations produced by the clustering of the galaxies, with the above assumption that their emissivities are equal,

$$N \sim x.$$

Therefore, if we denote by $\nu_n(b)$ the nth-order moments of the differential probability-distribution function for the number of observed galaxies at a latitude b, the Equations (10), applied to the fluctuations in the number of galaxies, will take the form

$$
\left.\begin{aligned}
\nu_1(b) &= a\bar{g}\, e^{-(1-\overline{q^{4.5}})\,\zeta\,\mathrm{cosec}\,b};\\
\nu_2(b) &= (a^2\bar{g}^2 + a\overline{g^2})\, e^{-(1-\overline{q^{3}})\,\zeta\,\mathrm{cosec}\,b};\\
\nu_3(b) &= (a^3\bar{g}^3 + 3a^2\overline{gg^2} + a\overline{g^3})\, e^{-(1-\overline{q^{4.5}})\,\zeta\,\mathrm{cosec}\,b};\\
&\dots\dots\dots\dots\dots\dots\dots\dots,
\end{aligned}\right. \qquad (11)
$$

where $a = \lambda R$ is the mathematical expectation of the number of clusters of galaxies up to the distance R, and $\bar{g}$, $\overline{g^2}$, etc., are the average number of galaxies in the cluster, the mean square number of galaxies in the clusters, etc.

Considering the counts of galaxies in grouped areas, characterized by various values of b, on the basis of each of the Equations (11) we can write a surplus system of equations. In view of the fact that the parameters characterizing the distribution of the galaxies in the Metagalaxy enter in the multipliers on the right-hand sides of the equations, while the parameters characterizing the properties of the dark matter in the exponents, the solution of the surplus system enables us to separate the two groups of parameters immediately.

Thus, the effect due to the tendency of galaxies to form clusters could be separated from the effect due to the nonuniform distribution of dark matter where there are fluctuations in the number of galaxies observed.

2. The Treatment of the Observational Material in the Catalogues of Hubble, Shapley and Shane and Wirtanen

After taking logarithms, Equations (11) have the form

$$
\left.\begin{aligned}
\log\nu_1(b) &= \log(a\bar{g}) - (1-\overline{q^{1.5}})\,\zeta\,\log e\,\mathrm{cosec}\,b,\\
\log\nu_2(b) &= \log(a^2\bar{g}^2 + a\overline{g^2}) - (1-\overline{q^{3}})\,\zeta\,\log e\,\mathrm{cosec}\,b,\\
&\dots\dots\dots\dots\dots\dots\dots\dots
\end{aligned}\right\} \qquad (12)
$$

In this form, they were used with the counts of the number of galaxies in areas of size 1° x 1° carried out by Hubble [6], Shapley [7] and Shane and Wirtanen [8]. For simplicity, in the following we will denote these counts by I, II and III, respectively.

The counts I were made on 760 areas, uniformly distributed over all the regions of the sky, for which $\delta > -30°$ and situated along circles of galactic latitude, for which b varied in steps of 5°. The limiting magnitude is 20^m. This enabled a surplus system of 18 equations to be written down for each of two first equations (12).

The counts II were made on 2175 areas in the region of the sky for which $\delta < +60°$. Here, the counts are qualified and the limiting magnitudes of the galaxies are different for different areas. The smallest limiting magnitude is $17^m.2$ and, therefore, this value was chosen as the limiting one in our calculations.

First of all, we introduced corrections for the distance of the areas from the center of the plate. For this, on all of the plates the total number of galaxies in an area of a given number was found, the ratio of this number to the number of galaxies in the central area was determined and then, in each plate, the number of galaxies in the given area was divided by the ratio obtained.

After this, the areas were grouped into strips of width 2° along the latitude and to each strip (and the areas included in this strip) was assigned the latitude of the center line of the strip. In all there were 29 strips with latitudes 1°, 3°, . . . 57°. This enabled a surplus system of 29 equations to be written down for each of the two first Equations (12).

The counts III consist of 6156 areas in the region of the sky: $[-20° < \delta < +20°,\ 12^h < \alpha < 18^h]$. The counts made in this work have not been corrected for a number of effects: atmospheric extinction, the effect of the distance of the area from the plate center, and others. However, corrections are given for the elimination of these effects in each area. After introducing these corrections we grouped the areas into strips of width 3° along the latitude and assigned to each strip (and the areas included in this strip) the latitude of the center line of the strip. In all there were 29 strips with latitudes 1.5°, 3.0°, . . ., 85.5°. This enabled a surplus system of 29 equations to be written down for each of the two first Equations (12).

In all of these systems of equations we discarded the equations for galactic latitudes equal to 3° or less, since in this case the value of cosec b was too large. After this, a weight equal to the number of galaxies counted in all of the corresponding strip was assigned to each of the equations and the surplus system were solved by the method of least squares.

The results are presented in the Table.

Author of the count	Hubble	Shapley	Shane and Wirtanen
	I	II	III
$2\log\overline{ag})$	$4.270+(\sigma_1=\pm0.34)$	$2.494+(\sigma_1=\pm0.096)$	$4.590+(\sigma_1=\pm0.070)$
$\lg(a^2\overline{g^2}+a\overline{g^2})$	$4.242+1.1\sigma_1\pm0.019$	$2.662+0.9\sigma_1+0.027$	$4.555+1.5\sigma_1\pm0,019$
a	—	2.18 ± 0.02	—
$\overline{g}$	—	8.3 ± 1.6	—
$(1-\overline{q^{1.5}})\zeta$	$0.394+(\sigma_2=\pm0.021)$	$0.272+(\sigma_2=\pm0.044)$	$0.918+(\sigma_2=\pm0.044)$
$(1-\overline{q^3})\zeta$	$0.601+2.2\sigma_2\pm0.023$	$0.433+1.8\sigma_2\pm0.023$	$1.604+3.0\sigma_2\pm0.023$
Δm	$0\overset{m}{.}46\pm0\overset{m}{.}02$	$0\overset{m}{.}3\,9\pm0\overset{m}{.}02$	$0\overset{m}{.}24\pm0\overset{m}{.}01$
ω	$0\overset{m}{.}39\pm0\overset{m}{.}09$	$0\overset{m}{.}2\,6\pm0\overset{m}{.}09$	$0\overset{m}{.}85\pm0\overset{m}{.}29$

In the calculation of the mean errors, as was to be expected, the residuals obtained in the equations for the systems, corresponding to the first and the second of Equations (12), are strongly correlated. The correlation coefficient r for the counts I, II and III was found to be 0.80, 0.91 and 0.95, respectively. Therefore, the sum of the squares of the residuals of the second system was in each case subdivided in the ratio $r:(1-r)$, its first part gave the error which is always proportional to the error of the first system and has the same sign (the coefficient of proportionality is in front of σ_1 (or σ_2)), while the second part gave the independent error.

Let us now investigate the results obtained. The comparison of the theoretical quantities in the first and second rows shows that the latter must be greater than the former. This predominance will be more marked the greater $\overline{g}$ and the smaller a.

The data of the Table show that, in the case of counts I and III, the quantity in the second row is not greater, but actually smaller than the quantity in the first row. However, as can be seen from the Table, this discrepancy is easily explained by chance errors. On the other hand, for the counts II the relation between the quantities in the first and second rows is correct and true to reality. Assuming that $\overline{g^2}\approx\overline{g}^2$, we have calculated the values of a and $\overline{g}$ for the counts II, which are given in the Table.

If we calculate for counts I and III the value of a, taking the above value of $\overline{g}$, then it is found to be equal to 16.4 and 23.8, respectively, i. e., considerably greater than the value of this quantity for the counts I. This explains why the addition of the term $a\overline{g^2}$ to the argument of the logarithm in the case of counts I and III did not play a significant role.

This explanation has the following physical meaning. In the counts II the limiting magnitude is small, galaxies from a small volume of the Metagalaxy are counted and the mean number of clusters of galaxies in an area is only 2.18. Therefore, chance fluctuations in the number of galaxies per area play an important role. On the other hand, in the case of counts I and III, the limiting magnitude and the volume of space observed is greater and the mean number of clusters per area is 16.4 and 23.8, rexpectively; therefore the chance fluctuations in the numbers of clusters do not play an important role and can be swamped by other chance phenomena. If the more common type of cluster in the Metagalaxy were the large clusters similar to the clusters in Virgo, Corona Borealis, and others, then in counts I and III also, the second row of the Table would have been greater than the first. In fact, small clusters of the type of the Local System predominate in the Metagalaxy. Here, it must be taken into account that in the counts of galaxies, dwarf galaxies must play a very insignificant role since such galaxies are observed only in our immediate vicinity. Therefore, the mean number of galaxies in a cluster, 8.3, must be understood in the sense of the mean number of normal and giant galaxies, without taking into account the dwarf ones. According to this idea, the average cluster of galaxies must be somewhat richer in composition than the Local System of galaxies.

The assertion that most of the galaxies are contained in clusters with a mean number of nondwarf galaxies equal to about 8 is the main result of our work.

It is interesting to point out that Neyman, Scott and Shane [4], on the basis of the counts III, assuming that the observed fluctuations in the number of galaxies are the result only of the tendency of the galaxies to form clusters, obtained for the mean number of galaxies in a cluster the result: $\bar{g} < 287$. From our work it becomes clear that this result is caused by the neglecting of the role of nonuniform absorption. If the nonuniform nature of the dark matter in the Galaxy is taken into account, then the results of the counts III, as we have seen, do not enable $\bar{g}$ to be determined.

The remaining results are given in rows 5-8. Assuming that $q^3 \approx (\overline{q^{1\cdot 5}})^2 = \overline{q^3}$, $\bar{q}$ and $\overline{\Delta m} \approx 1\cdot 08 e^{-\bar{q}}$, the mean absorption of one cloud, was calculated in stellar magnitudes. The latter was found to be, on the average, somewhat higher than the values obtained earlier by V. A. Ambartsumian and some other authors. The fact that we found $\overline{\Delta m}$ for the counts III to be considerably lower than for the counts I, can be explained by assuming that the system of corrections, proposed by the authors of the counts III, was based on taking into account the systematic variation in the number of galaxies between different areas and different plates. These corrections, in addition to the elimination of true systematic variations, must also lead to some smoothing of the chance fluctuations (i. e., a certain part of the inherent fluctuations will contribute to the quantity which will be considered as a systematic variation) and lead to a decrease of $\overline{\Delta m}$. In the counts II the corrections were introduced to take into account only the systematic variation between different areas and, therefore, the smoothing effect must play a smaller role here.

The mean value of the absorption in the direction of the galactic pole was found from the equation

$$\omega = \zeta \overline{\Delta m}.$$

Here, we see that there are considerable discrepancies among the data, particularly those obtained from counts III, on the one hand, and I and II on the other. Apparently this discrepancy is explained by the fact that the region of the sky in which the counts III were made is actually richer in dark matter than the other regions.

A. A. Zhdanov Leningrad State University Received December 6, 1956

LITERATURE CITED

[1] V. A. Ambartsumian, Biull. Abastum. Obs. 4, 17 (1940).

[2] V. A. Ambartsumian, Soobshch. Biurak. Obs. 6 (1951).

[3] J. Neyman and E. L. Scott, Astrophys. J. 116, 144 (1952).

[4] J. Neyman, E. L. Scott and C. D. Shane, Astrophys. J. 117, 92 (1953).

[5] T. A. Agekian, Vestnik Leningrad Univ., No. 1, Ser. 1, 145 (1956).

[6] E. Hubble, Astrophys. J. 79, 8 (1934).

[7] H. Shapley, Harvard Annals 105, 137 (1937).

[8] C. D. Shane and C. A. Wirtanen, Astron. J. 59, 285 (1954).

THE DISTRIBUTION OF MATTER IN THE POLAR RAYS OF THE SOLAR CORONA

N. I. Dziubenko

A photometric investigation of 9 polar rays of the solar corona was made from photographs obtained at Kozeltze on June 30, 1954 with the 10 meter coronagraph of the expedition of the Astronomy Department of Kiev University.

The variation of electron concentration in the polar ray as a function of the distance from the center of the sun was derived, assuming that the ray is of uniform density in a cross section (circular) and lies in the plane of projection. It is represented by the equation: $N_e = N_0 r^{-n}$. The parameter n for the rays studied varies within the limits 6 to 11.

The ratio of the electron concentrations in the ray and inter-ray space (N_r / N_c) and its variation with height were determined. The Fraunhofer component of coronal emission was accounted for using the values of $I_c/(I_c + I_F)$ found by comparing the brightness of the polar corona of 1954 with that of the Fraunhofer corona according to van de Hulst. For the bright rays under consideration on the average, $N_r / N_c = 9$. For most of the rays N_r/N_c is constant within the limits of error or increases with height.

The attempt to represent the polar rays as a steam of corpuscles rapidly receding from the sun meets with great difficulties. The diffusion of plasma along lines of force of the solar magnetic field is a more acceptable description. Assuming that the density gradient in the rays is determined only by the kinetic temperature and gravity, the temperature of the rays has been found. The derived value is too large. This shows that there are other factors, besides the two mentioned above, which determine the density gradient in the polar rays.

1. For the study of the distribution of matter in the polar rays of the solar corona, a photograph of the corona was used which was taken on June 30, 1954 by the expedition of the Kiev University Department of Astronomy with a 10-meter horizontal camera (S. K. Vsekhsviatskii, G. K. Nazarchuk, V. I. Ivanchuk) [1]. The intermediate and inner corona were obtained on a plate Agfa Astro 30 × 40 cm^2 ($\lambda_{eff} \approx 4300$ A) with an exposure of 70 sec (see the figure in Reference [2]).

The negative of the corona was measured with a microphotometer MF-4 of the Kiev Astronomical Observatory. Nine brightest rays were chosen for photometric measurements: 5 in the northern and 4 in the southern polar region (Fig. 1).

Ray No. 1. The ray is bright, and can be traced on the photograph to a distance of 1 $R_\odot$ from the sun's limb. The width of the ray increases with distance from the limb. The eastern edge of the ray is sharper than the western.

Ray No. 2. It is bright and sharp; it can be followed to a distance of 1 $R_{\odot}$ from the limb. It widens quickly with height. The edges of the ray are sharp.

Ray No. 3. It is the eastern ray of a close doublet. It can be followed to a distance of 1 $R_{\odot}$ from the limb and almost does not widen with height. The eastern edge of the ray is considerably sharper than the western.

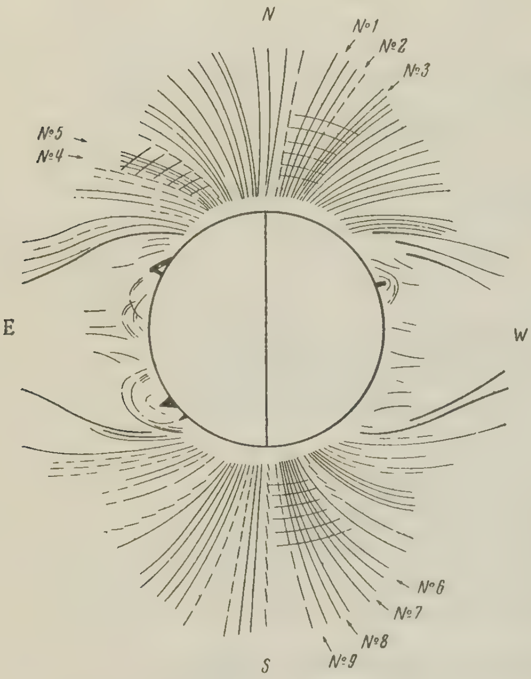


Fig. 1. The location of the rays studied. The direction of the photometric sections are indicated by concentric arcs.

Ray No. 4. It can be followed approximately to 1 $R_{\odot}$ from the limb. It widens little with height. At large distances from the limb doubling is visible. It is possible that it consists of two superimposed rays which lie near the plane of the photograph. The edges are not very sharp.

Ray No. 5. The ray is thin, single and widens with height to a lesser extent than the rays considered earlier and can be followed to a smaller distance from the limb. At a distance of 0.5 $R_{\odot}$ a doubling is visible which is reminiscent of a shift. The edges are sufficiently sharp.

Rays No. 6 and 7. This is a typical doublet. The rays are very bright and can be followed to a distance of 1 $R_{\odot}$ from the limb. The western ray (No. 6) is considerably wider than the eastern (No. 7). Ray No. 6 widens much with height, while ray No. 7 does not visibly change its width. The eastern edge of the doublet is sharper than the western. The inner boundaries of the rays are sufficiently sharp. The brightness of ray No. 6 decreases with height more rapidly than that of ray No. 7.

Rays No. 8 and 9. A bright typical doublet. It can be followed to 1 $R_{\odot}$ from the limb. The western ray (No. 8) is narrower than the eastern (No. 9) and is approximately of the same brightness. The western edge of the doublet is sharper than the eastern, which is diffuse. The inner edges of the doublet are sharp. Ray No. 8 widens very little with height, ray No. 9 widens considerably more.

The rays were measured photometrically in a transverse direction along concentric arcs parallel to the solar limb. For this purpose a template was used which was made by drawing concentric circles on a glass plate. The microphotometer readings were taken visually.

For each ray 17-18 sections were measured for $1.23 < r < 1.76$ and a graph of the blackening was constructed for each photometric section. For example, several such graphs for rays Nos. 6, 7, 8 and 9 are given in Figure 2.

With the help of a characteristic curve the degree of blackening, corresponding to the ray and the corona background obtained for the region of a ray by interpolation, was reduced to relative intensities (I_{r+b} and I_b, respectively).

The "background" consisted of the blackening produced by the "diffuse" corona, the Fraunhofer component of the coronal emission, the scattering of light in the Earth's atmosphere, in the instrument and the emulsion layer. The brightness of a ray (I_r) was taken to be its excess over the "background". The value of I_r for a given section was obtained from $I_{r+b} - I_b$. The values of I_r smoothed graphically (see Figure 3 for an example) and expressed in arbitrary units of brightness are given in Table 1.

The relative values of the electron concentration in the rays were determined from the equation:

$$N_e(r) = \frac{I_r(r)}{\Delta(r)\cdot W(r)}, \tag{1}$$

where $\Delta(r)$ is the width of the ray at a distance $\underline{r}$ from the center of the sun and $W(r)$ is the dilution coefficient.*

TABLE 1

r in $R_\odot$ \ Ray	1	2	3	4	5	6	7	8	9
1.250	152	295	147	132	88	303	220	228	207
1.300	87	147	87	84	66	199	151	121	115
1.350	55	86	56	59	48	112	100	63	77
1.400	37	56	39	42	33	62	66	38	51
1.450	25,0	36	28.5	30	21.0	40	43	26.5	33
1.500	18.0	25.0	22.0	22.5	14.0	26.5	27.0	19.5	23.0
1.550	14.5	17.5	18.0	17.0	10.0	18.5	18.5	15.0	17.5
1.600	11.0	13.5	14.5	12.0	7.5	13.5	14.5	11.5	13.0
1.650	8.5	10.0	11.5	9.0	5.5	10.5	11.5	8.5	9.5
1.700	6.0	6.5	9.0	6.5	4.0	8.0	9.0	6.5	6.5
1.750	—	—	—	4.0	3.0	6.0	7.0	5.0	4.0

TABLE 2

r in $R_\odot$ \ Ray	1	2	3	4	5	6	7	8	9
1.250	0.044	0.053	0.051	0.029	0.021	0.046	0.033	0.028	0.033
1.300	0.047	0.055	0.053	0.030	0.027	0.050	0.034	0.030	0,036
1.350	0.049	0.058	0.055	0.031	0.032	0.055	0.035	0.033	0.039
1.400	0.053	0.060	0,055	0.032	0.035	0.060	0.036	0.035	0.041
1.450	0.055	0.063	0.055	0.033	0.037	0,066	0.037	0.037	0.043
1.500	0.057	0,066	0.056	0.034	0.039	0.070	0.038	0.039	0.045
1.550	0.060	0.069	0.056	0.035	0.040	0.075	0.039	0·041	0.048
1.600	0.063	0.071	0.056	0.036	0.042	0.080	0.040	0.042	0.051
1.650	0.065	0.073	0.056	0.038	0.042	0.084	0.042	0.044	0.053
1.700	0.068	0.076	0.056	0.039	0.042	0.088	0.043	0.045	0.054
1.750	—	—	—	0,040	0.043	0,092	0.043	0.047	0.056

The width of the rays was taken as the half-width of the brightness curve along a transverse section of the ray (see Reference [3]).

The values of Δ in units of $R_\odot$ are given in Table 2.

The dilution coefficient for photographic light was taken from Reference [4].

The relative values of the electron concentration in the rays are given in Table 3.

The variation of the electron concentration in the rays is represented well by the equation:

$$N_e(r) = N_0 r^{-n}. \tag{2}$$

The parameter $\underline{n}$ was determined by the least-squares method (Table 4).

How well the analytical Expression (2) describes the actual decrease of the mean electron concentration in the ray can be seen from Figure 4.

Since for ray No. 8 it is impossible to describe the decrease of N_e for $1.250 < r < 1.750$ by a single power law with constant $\underline{n}$ (Fig. 4), two values of this parameter were determined:

---

*Equation (1) is based on the following assumptions: 1) the ray lies in the plane of the photograph, 2) the transverse cross section of the ray is a circle and 3) the electron concentration in the transverse sections of the ray is constant.

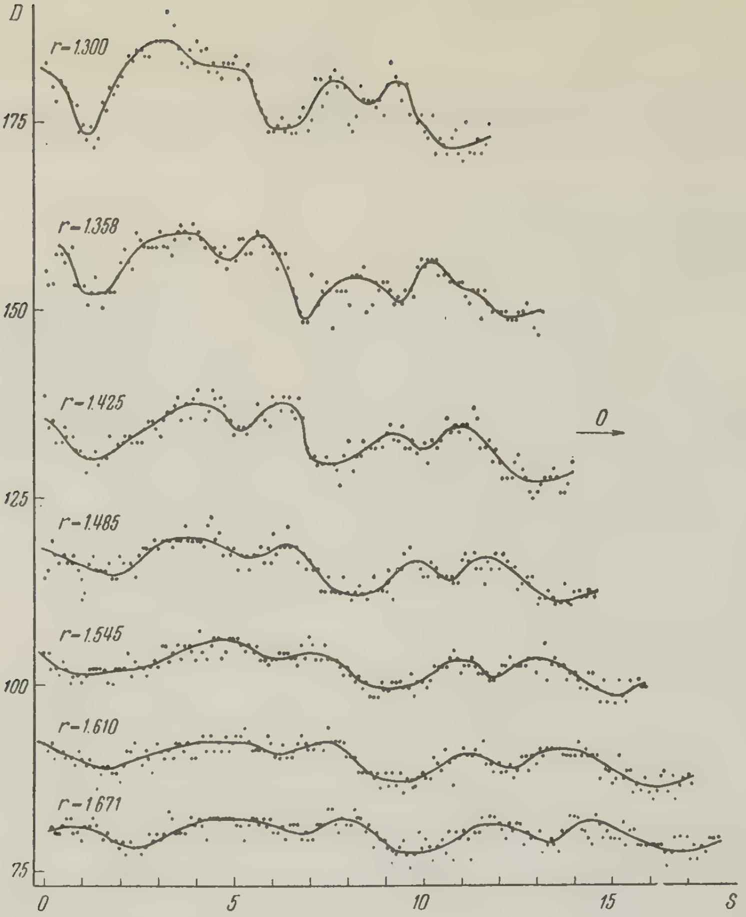


Fig. 2. The curves of the blackening along the photometric sections of the rays Nos. 6, 7, 8 and 9.

TABLE 3

r in $R_\odot$ \ Ray	1	2	3	4	5	6	7	8	9
1.250	236	385	198	312	287	452	456	558	430
1.300	139	199	123	210	184	299	334	304	240
1.350	92.7	123	84.2	157	124	168	236	158	163
1.400	63.5	86.5	64.5	119	85.8	95.5	167	98.7	113
1.450	46.0	57.7	52.4	91.8	57.4	61.2	117	72.4	77.6
1.500	34.7	42.2	43.2	72.8	39.4	41.6	78.1	54.9	56.2
1.550	29.1	31.0	38.8	58.5	30.1	29.8	57.2	44.2	44.0
1.600	22.7	24.5	33.7	43.3	23.2	21.9	47.1	35.6	33.1
1.650	18.5	19.0	28.5	32.9	18.2	17.4	38.1	26.8	24.9
1.700	13.2	12.8	24.0	24.9	14.2	13.6	31.2	21.6	18.0
1.750				15.9	11.1	10.3	25.8	16.9	11.3

TABLE 4

Ray	1	2	3	4	5	6	7	8	9
n	8.83	10.1	6.09	7.64	9.22	11.1	8.68	9.66	9.94

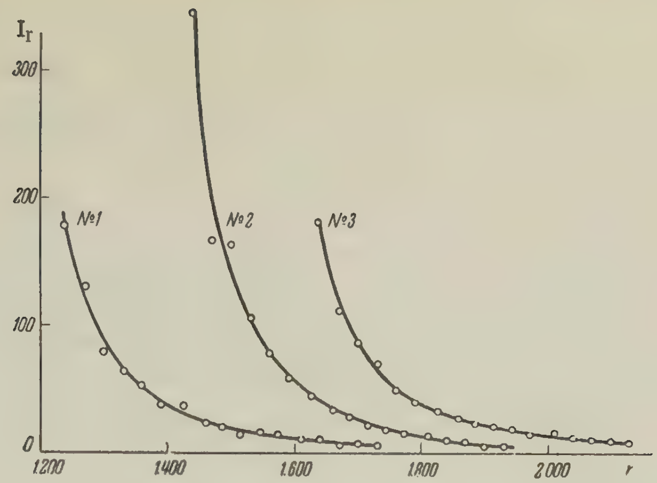


Fig. 3. The curves for the rays Nos. 2 and 3 have been shifted in the positive direction along the abscissa by 0.200 and 0.400, rexpectively.

$$n_1 = 15.2 \ (1.250 < r < 1.400),$$
$$n_2 = 6.19 \ (1.400 < r < 1.750).$$

In Table 4 for ray No. 8 an approximate value of n for the interval $1.250 < r < 1.750$ is given.

2. To determine the ratio of the electron concentrations in the ray and in the space between the rays the following equation, applied in Reference [3], was used:

$$\frac{N_r}{N_C} = 1 + \sqrt{\pi}\,\frac{\Gamma(n/2)}{\Gamma((n+1)/2)}\,\frac{\rho}{\Delta}\,\frac{I_r}{I(\rho)}. \qquad (3)$$

The Fraunhofer component of the coronal emission was taken into account in the following way. Using the determination of the integral luminosity of the corona on June 30, 1954, the absolute brightness corresponding to the isophotes of the corona could be found. For this purpose an original drawing of the isophotes of the corona in 1954, made by Iu. A. Nadubovich from a photograph taken with the red coronagraph of the Kiev University Department of Astronomy [1], was used. The features of the isophotes considered were given in Reference [5].

Iu. A. Nadubovich has determined the law for the decrease of the surface brightness of the corona on June 30, 1954:

$$I(\rho) = \begin{cases} 1000\,\rho^{-8.76} \ (1.11 < \rho < 1.30), \\ 123\,\rho^{-5.15} \ (1.30 < \rho < 1.82). \end{cases} \qquad (4)$$

Let the coefficient for the transformation from the arbitrary unit of brightness, adopted here, to the unit B (10^{-6} is the brightness of the center of the solar disc) be equal to p. Then the integral luminosity of the corona relative to the luminosity of the sun is determined from the equation:

$$\frac{E_C}{E_\odot} = p\,\frac{2\int_1^\infty I(\rho)\,\rho\,d\rho}{I_\odot}, \qquad (5)$$

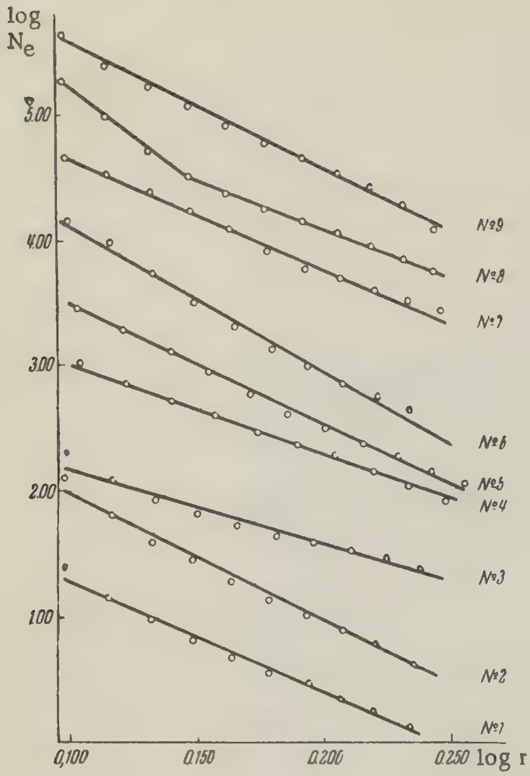


Fig. 4. The electron concentration in the polar rays. The graphs have been spread out along the ordinate at intervals of 0.50.

where $I(\rho)$ is taken in the units adopted in (4) and $\bar{I}_{\odot}$ is the mean brightness of the solar disc in B units.

In photographic light (coefficient of limb darkening $u = 0.8$), as is easily shown,

$$\bar{I}_{\odot} = 0.733 \cdot 10^{6}.$$

Extending the law (4) to the whole of the interval of distances from 1 $R_{\odot}$ to ∞ (in extrapolating to the limb we take the first expression, in extrapolating to ∞ — the second) and using Equation (5), we find

$$\frac{E_c}{E_{\odot}} = 3.82 \cdot 10^{-4}\, p. \qquad (6)$$

On the other hand, Table 5 gives determinations of the integral luminosity of the corona on June 30, 1954 in the photographic region [6].

The weighted mean is

$$\frac{E_c}{E_{\odot}} = 0.55 \cdot 10^{-6}. \qquad (7)$$

Comparing (6) and (7), we obtain

$$p = 1.44 \cdot 10^{-3}.$$

Knowing p, it is easy to find the absolute brightness corresponding to an isophote in the corona from the mean radius of the isophote by using (4). In our case the brightness of the isophote 6 (the isophote numbers are taken according to Reference [5]) was determined in this way. The brightness corresponding to the other isophotes was determined in relation to the isophote 6 with the help of the known intervals between the isophotes in stellar magnitudes (see Table in Reference [5]).

TABLE 5

Author	Camera	$\frac{E_c}{E_{\odot}}$	Weight
A. P. Stefanov P. Ia. Sukhoivanenko	Without an objective	$0.30 \cdot 10^{-6}$	0.5*
N. N. Sytinskaia	Without an objective	$0.62 \cdot 10^{-6}$	1
N. S. Orlova	Without an objective	$0.60 \cdot 10^{-6}$	1

---

*The result of A. P. Stefanov and P. Ia. Sukhoivanenko is taken with a weight 0.5 because during the determination the difference in the transparency of the atmosphere at the time of the eclipse and at the time of calibration using the Moon was not taken into account.

The absolute brightness corresponding to the isophotes of the corona in 1954, determined in this way, are given in Table 6 (I).

TABLE 6

Isophot	r_p	I (Bunits $\cdot 10^{-2}$)	I_F (Bunits$\cdot 10^{-2}$)	$\frac{I_c}{I_c+I_F}$
3	1.100	69.0	11.6	0.833
4	1.125	50.5	10.3	0.796
5	1.145	39.1	9.4	0.760
6	1.175	29.4	8.2	0.721
7	1.205	22.3	7.1	0.682
8	1.250	16.4	5.95	0.635
9	1.300	11.9	4.97	0.582
10	1.345	9.02	4.30	0.524
11	1.405	6.23	3.53	0.433
12	1.495	4.30	2.70	0.372
13	1.595	3.12	2.13	0.317
14	1.730	1.93	1.59	0.176

In the second column of Table 6 are given the mean radii of the isophotes in the polar regions of the corona (r_p). The averaging is carried out over sectors, symmetrically situated relative to the solar axis, of the southern and northern polar corona with opening angles of 60 heliocentric degrees.

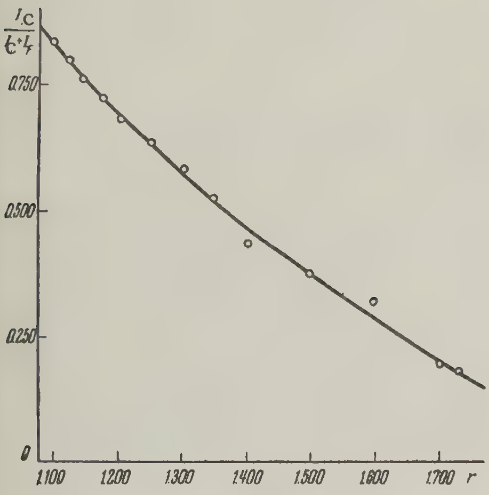


Fig. 5.

The brightness of the Fraunhofer component (I_F) is taken according to van de Hulst [7]. Let us point out that the Fraunhofer component of the coronal emission according to the model of van de Hulst is in good agreement with the recently-published results of Elsässer [8], who has calculated the distribution of brightness of the F corona on the basis of data obtained from the observation of the brightness of Zodiacal light.

The values of the ratio $I_c/(I_c + I_F)$ $(I = I_c + I_F)$ given in Table 6 have been averaged graphically (Fig. 5). The final result is given in Table 7, where for comparison the values of $I_c/(I_c + I_F)$ for the van de Hulst model of the polar corona are also given.

The ratio of the electron concentrations N_r/N_c have been determined from Equation (3) taking into account the Fraunhofer component using the value of the ratio $I_c/(I_c + I_F)$ and, for comparison, without taking into account the F corona. For rays Nos. 6 and 7, the ratio N_r/N_c was also determined with the Fraunhofer component taken into account using the values of $I_c/(I_c + I_F)$ taken from van de Hulst [7].

The values of $\underline{n}$ (see Equation (3)) for the "diffuse corona" near the rays considered were determined approximately from the graphs of the decrease of the surface brightness in a logarithmic plot, log I against log ρ (it is to be remembered that for Equation (3) it is sufficient to know approximate values of $\underline{n}$). The values of $\underline{n}$ determined in this way, with and without taking into account the F corona, as well as the values of the quantity

$$k = \sqrt{\pi}\,\frac{\Gamma(n/2)}{\Gamma((n+1)/2)}$$ are given in Table 8.

The values of I_r and Δ were taken from Tables 1 and 2.

In Table 9 are presented the ratios N_r/N_c for the rays investigated, with and without taking into account the Fraunhofer component. N_r/N_c and N_r/N'_c are the ratios of the electron concentrations with the Fraunhofer component taken into account and not taken into account, respectively.

TABLE 7

ρ in $R_{\odot}$	1.250	1.300	1.350	1.400	1.450	1.500	1.550	1.600	1.650	1.700	1.750
$I_c/(I_c+I_F)$ 1954	0.634	0.575	0.520	0.467	0.415	0.368	0.323	0.280	0.237	0.198	0.158
$I_c/(I_c+I_F)$ [van de Hulst]	0.539	0.459	0.395	0.338	0.294	0.257	0.228	0.204	0.182	0.164	0.146

TABLE 8

Ray	№ 1, 2, 3,		№ 4, 5		№ 6, 7		№ 8, 9	
	n	k	n	k	n	k	n	k
with I_F	12.2	0.730	12.2	0.730	12.4	0.725	12.7	0.716
without I_F	8.63	0.876	8.47	0.885	8.90	0.863	8.67	0.874

We see that if the F component of the coronal emission is not taken into account, this leads to a significant distortion of the actual values of N_r/N_c and changes the behavior of this ratio as a function of the distance from the edge of the solar disc.

Undoubtedly, the method used here for taking into account the Fraunhofer component of the coronal emission is accompanied by a considerably error (an uncertainty of the order of several tens of per cent at large ρ). The results obtained, however, are the first approximation to reality and enable some conclusions to be drawn.

It must be pointed out that the use of corona isophotes which were obtained from the photographs made with the red coronagraph ($\lambda_{eff} \approx 6400$ A), with all the other quantities and calculations referring to the photographic region ($\lambda_{eff} \approx 4300$ A), cannot by itself give a large error. The fact is that with the same distribution of energy in the continuous spectra of the corona and of the center of the sun (this can be considered verified to a high degree of accuracy for the region of the corona studied by us) the brightness of any point in the corona, expressed in B units, does not depend on the spectral region used for observation, since the brightness unit itself is usually always referred to the same spectral region. It is clear that in these conditions the corona isophotes constructed for different spectral regions will be identical, both in shape and in the value of the brightness, expressed in B units, corresponding to them.

The variation of N_r/N_c with distance from the center of the solar disc for the nine rays investigated is presented graphically in Figure 6.

3. The results obtained above enable some preliminary ideas to be put forward concerning the mechanism for the transport of matter in the polar rays of the solar corona.

Before we proceed to the discussion of these results, let us point out the following:

1) The features of the structure of the solar corona clearly indicate the presence of a general magnetic field of the sun, extending to a considerable distance from the sun's surface.

2) All the transport processes (for example, diffusion, thermal conduction) in a sufficiently rarefied and ionized medium can mainly take place in the direction of the magnetic-force lines [4]. Even a very weak magnetic field practically eliminates the possibility for diffusion and thermal conduction in a direction perpendicular to the magnetic-force lines.

If we consider polar rays to be flows of matter along the force lines of a general magnetic field of the sun, then, neglecting diffusion through the "walls" of a ray (see Paragraph 2, above), we can apply the continuity equation to such a flow [7]:

TABLE 9

r	1.250	1.300	1.350	1.400	1.450	1.500	1.550	1.600	1.650	1.700	1.750
Ray № 1											
N_r/N_c	5.1	5.0	5.1	5.2	5.4	5.7	6.4	6.8	7.5	7.7	
N_r/N'_c	4.1	3.7	3.6	3.3	3.2	3.1	3.1	2.9	2.8	2.6	
Ray № 2											
N_r/N_c	7.6	6.7	6.5	6.6	6.6	6.6	6.7	7.3	7.8	7.5	
N_r/N'_c	6.0	4.9	4,4	4.1	3.8	3.5	3.2	3.1	2.9	2.6	
Ray № 3											
N_r/N_c	4.5	4.6	4.8	5.4	6.2	7.2	8.7	10.1	11.8	13.8	
N_r/N'_c	3.7	3.5	3.4	3.5	3.6	3.7	4.0	4.0	4.1	4.0	
Ray № 4											
N_r/N_c	6.6	7.0	8.1	9.0	10.0	11.0	12.0	11.9	12.5	13.6	12.4
N_r/N'_c	5.3	5.2	5.5	5.6	5.5	5.4	5.3	4.7	4.3	4.0	3.2
Ray № 5											
N_r/N_c	5.7	5.9	6.4	6.7	6.4	6.3	6.5	7.0	7.3	8.0	8.7
N_r/N'_c	4.6	4.4	4.4	4.2	3.7	3.3	3.2	3.0	2.8	2.7	2.5
Ray № 6											
N_r/N_c	9.2	9.9	8.9	7.6	7.2	7.1	7,1	7.2	7.9	8.6	9.4
N_r/N'_c	7.2	7.1	5.9	4.6	4.1	3.6	3.3	3.1	3.0	2.8	2.6
N_r/N''_c	10.6	12.2	10.4	10.0	9.7	9.6	9.5	9.5	10.0	10.0	10.0
Ray № 7											
N_r/N_c	9,3	11.0	12.1	12.4	12.8	12.3	12.6	14.3	16.2	18.4	22.0
N_r/N'_c	7.3	7.8	7.9	7.4	6.8	6.0	5.5	5.4	5.3	5.1	5.0
N_r/N''_c	10.7	13.4	15.5	16.6	17.4	17.0	17.3	19.2	20.7	21.8	23.6
Ray № 8											
N_r/N_c	11.4	10.2	8.6	8.1	8.5	9.2	10.0	11.2	12,0	13.5	15.0
N_r/N'_c	9.1	7.5	5.8	5.0	4.8	4.7	4.6	4.5	4.2	4.0	3.7
Ray № 9											
N_r/N_c	9.3	8.3	8.9	9.1	9,1	9.5	10.1	10.3	11.0	11.2	10.5
N_r/N'_c	7.4	6.2	6.0	5.6	5.1	4.8	4.6	4.2	3.9	3.4	2.8

N_r/N''_c is the ratio of the electron concentrations with the Fraunhofer component taken into account according to van de Hulst.

$$v(r)\,\Delta^2 N_e = \text{const}. \tag{8}$$

Table 10 gives the values of the ratio

$$d = \frac{(\Delta^2 N_e)_{r=1.250}}{(\Delta^2 N_e)_{r=1.700}}$$

for the rays investigated by us.

TABLE 10

Ray	№ 1	№ 2	№ 3	№ 4	№ 5	№ 6	№ 7	№ 8	№ 9
d	7.5	14.6	6.9	6.9	5.0	9.1	8.6	10.2	8,9

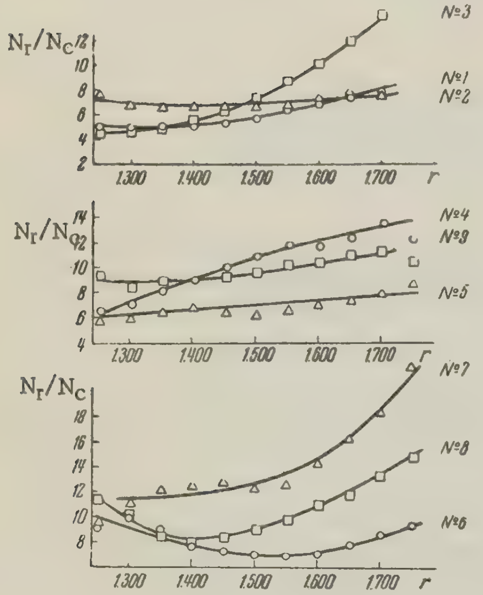


Fig. 6.

Therefore, the product $\Delta^2 N_e$ in the interval from $r = 1.250$ to $r = 1.700$ decreases by 5-10 times. According to Equation (8), in the same interval of r, the velocity of the matter in the ray must increase by the same factor. It follows from this that a definite mechanism for the acceleration of the particles in the beam must be in operation. At the present time, the most probable is the acceleration of particles in the varying electromagnetic fields of the disturbed regions of the solar surface.

However, here we encounter great difficulties. The fact is that for equatorial rays of the solar corona, which undoubtedly represent flows of matter, the quantity $d \approx 4$ (according to the results of the diploma work of Iu. A. Nadubovich). Therefore, in the interval from $r = 1.250$ to $r = 1.700$ the velocity of the particles grows by a factor of 4, i. e., approximately 2 times less than for the polar rays according to our observations. Yet it is hard to imagine that the electromagnetic mechanism would be twice as effective in the polar regions of the sun than in the equatorial regions.

This difficulty can be partially overcome if we consider the polar rays, in contrast to the equatorial rays, not as flows of fast corpuscles, but as formations occurring as the result of the diffusion of coronal matter along the force lines of the general magnetic field of the sun [4]. According to the calculations of I. S. Shklovskii, the time for diffusion equilibrium to be established in the corona is of the order of several hours. After this period ot time, there will be established in the ray a density gradient which depends only on the strength of the solar gravitation and the kinetic temperature of the matter. On the other hand, the rate of diffusion in a direction perpendicular to the magnetic-force lines will be negligibly small. This guarantees the stability of the ray for a considerable period of time, even if the magnetic-field intensity is very small.

4. Let us evaluate the kinetic temperature in the polar rays from the known density gradient, using the equation for hydrostatic equilibrium (for the applicability of the latter to the solar corona see Reference [9])

$$\frac{dP}{dh} = -g\rho, \tag{9}$$

where h is the height and ρ is the density of matter.

From this, considering that the coronal plasma consists only of protons and electrons ($P = 2N_e kT$; $\rho = N_e m_H$), that the kinetic temperature for the heights considered is constant (T = const) and taking the relation $g = GM_\odot / r^2$ into account, it is easy to find that

$$N_e = Ce^{\alpha/r}, \tag{10}$$

where $\alpha = GM_\odot m_H / 2kTR_\odot$, C is a constant and r is expressed in terms of $R_\odot$. Therefore,

$$T = \frac{GM_\odot m_H}{2k\alpha R_\odot}. \tag{11}$$

It is not difficult to show that the quantity $N_e = N_0 r^{-n}$ for the interval $1.250 < r < 1.700$ can be also represented with a sufficient accuracy by an equation of the form (10), where

$$\alpha = \frac{n}{0.684}. \tag{12}$$

Substituting (12) into (11), we obtain

$$T = \frac{0.684\, GM_\odot m_H}{2knR_\odot}. \tag{13}$$

The mean kinetic temperatures of the matter in the rays, determined from Equation (13) (n is taken from Table 4), are given in Table 11.

TABLE 11

Ray	1	2	3	4	5	6	7	8	9
$T \cdot 10^{-6}$	0.90	0.78	1.30	1.03	0.86	0.71	0.91	0.82	0.80

At the temperatures in the rays obtained here, the emission in the green coronal line must predominate over the emission in the red line [10]. However, from observations when there are no eclipses it is known that in the time of the minimum of solar activity the green line is almost never observed in the polar regions of the corona. This forces the conclusion that the density gradient in the rays is determined not only by the kinetic temperature and the force of gravity, but also by other factors. Neglecting the latter resulted in values of the kinetic temperature which are too high.

In conclusion, the author expresses his deep gratitude to Professor S. K. Vsekhsviatskii and G. M. Nikol'skii for valuable advice and directions.

Department of Astronomy, Kiev State University.

Received September 7, 1956

LITERATURE CITED

[1] S. K. Vsekhsviatskii, Astr. Tsirk. AN SSSR 1954 No. 151.

[2] S. K. Vsekhsviatskii and G. M. Nikol'skii, Astron. J. (USSR) 32, 354 (1955).

[3] G. M. Nikol'skii, Astron. J. (USSR) 33, 87 (1956).

[4] I. S. Shklovskii, The Solar Corona [in Russian] (GITTL, 1951).

[5] Iu. A. Nadubovich, Astr. Tsirk AN SSSR 1955 No. 161.

[6] A. P. Stefanov and P. Ia. Sukhoivanenko, Astr. Tsirk. AN SSSR 1955, No. 156; N. N. Sytinskaia, Astr. Tsirk. AN SSSR 1955, No. 161; N. S. Orlova, Astr. Tsirk. AN SSSR 1955 No. 165.

[7] H. C. van de Hulst, Bull. Astron. Netherlands, 11, No. 410, 135 (1950).

[8] H. Elsässer, Z. Astrophys. 37, 114 (1955).

[9] S. B. Pikel'ner, Izv. Krymsk. Astrofiz. Obs. 5, 34 (1950).

[10] M. Waldmeier, Ergebnisse und Probleme der Sonnenforschung (Leipzig, 1955).

ON ONE CLASSIFICATION OF THE CYCLES OF SUNSPOTS AND FACULAE

J. Xanthakis

The article deals with the study of properties of solar cycles, but here, contrary to previous practice, the author, in dealing with both the sunspots and the faculae, makes use of the curves obtained directly from observations and not of the smoothed-out curves. The cycles are divided into four groups according to the value of the ordinate of the relative number of sunspots at their maximum. The article indicates that the secondary fluctuations normally occur in cyclic curves at approximately the same moments before and after the moment of maximum. From this the author concludes that each solar cycle exhibits its specific mechanism, and that fluctuations in the cycles of the same group follow a more or less regular rhythm.

It is generally known that the curves depicting the change in the number of sunspots differ from one another in their ordinates at the moments of maximum and in the time intervals dividing the moments of maximum and minimum.

According to M. Waldmeier, every curve is composed of isolated fluctuations — "flare-ups". Each flare-up starts at a certain moment, reaches a maximum of intensity, and then diminishes asymptotically toward zero. Thus, according to Khalm's* hypothesis, the sunspot curve represents a summation of consecutive, overlapping flare-ups, and the observable minimum of activity corresponds to the overlap of the declining phase of the preceding cycle and the ascending phase of the next cycle. Waldmeier concluded from the foregoing analysis that the consecutive sunspot cycles behave in different ways. Little time is needed to diminish the number of sunspots in the cycles with a high maximum, while the diminution of this number in a cycle of a relatively low maximum requires more time. He determined also the fact that the curves of sunspots constitute monoparametric families of curves, with the parameter of each family represented by the ordinate R_m which the curve reaches at the moment of maximum; the location of this ordinate varies with the period of ascension T, with the period of decline, and also with other characteristic factors.

M. Waldmeier arrived at these conclusions from the reduced values of the relative sunspot numbers, R. We know, however, that the difference between the relative sunspot number obtained by observations and the number derived by reducing the observations (smoothing out the values) is significant, and sometimes very large. For this reason, a study of the sunspot and faculae curves obtained directly from observation and without having the relative values reduced, is of some interest. It is with such a study that this article is concerned, and the early results of this study are presented below.

Table 1 contains the dates of the start, the maximum and the minimum for each cycle, and also the corresponding values of R_m (the observed maximum of relative numbers), the period of ascension T, and the period of decline T'. The last colume of the table contains the values of the ratio $q = T/T'$.

---

* Transliteration of Russian — Publisher's note.

TABLE 1

Cycles								
No.	Date	Date of start	Date of max.	Date of min.	R_m	T(month)	T'(month)	$q = \frac{T}{T'}$
1	1755—66	1755 June	1761 May	1766 June	107.2	72	61	1.180
2	1766—75	1766 July	1769 Oct.	1775 Feb	158.2	40	64	0.625
3	1775—84	1775 March	1778 May	1784 May	238.9	39	72	0.542
4	1784—98	1784 June	1787 Nov.	1798 June	174.0	43	125	0.344
5	1798—1810	1798 July	1804 Oct.	1810 June	62.3	77	68	1.132
6	1810—23	1810 July	1817 May	1823 Aug.	96.2	81	77	1.052
7	1823—33	1823 Sept.	1830 Apr.	1833 June	107.1	80	38	2.105
8	1833—43	1833 July	1836 Nov.	1844 June	206.2	42	90	0.467
9	1843—56	1844 July	1847 Oct.	1856 May	180.4	40	98	0.500
10	1856—67	1856 June	1860 June	1867 Jan.	116.7	50	78	0.641
11	1867—78	1867 Feb.	1870 May	1878 Aug.	176.0	40	99	0.444
12	1878—89	1878 Sept.	1884 Jan.	1889 Jan.	91.5	65	60	1.083
13	1889—1901	1889 Feb.	1893 Aug.	1901 Apr.	129.2	55	92	0.598
14	1901—13	1901 May	1907 Feb.	1913 May	108.2	70	75	0.933
15	1913—23	1913 June	1917 Aug.	1923 Aug.	154.5	51	72	0.709
16	1923—33	1923 Sept.	1928 July	1933 Dec.	98.0	59	65	0.908
17	1933—44	1934 Jan.	1937 July	1944 Jan.	145.1	43	81	0.531
18	1944—54	1944 May	1947 May	1954 Jan.	201.3	36	68	0.529

We should note first of all that in our case, as well as in the case of the reduced values, there exists a close correlation between the observed value of R_m and the period of the ascension T, but that such a correlation does not exist between R_m and the period of decrease T'. These facts are shown in Figures 1 and 2.

On the basis of this observation we divided the sunspot cycles into four groups according to the periods of asension T. Group I includes cycles 2, 3, 4, 8, 9, 11, 17 and 18, for which the period of ascension varies between the limits of 35 to 45 months. Group II contains cycles 10, 13 and 15, with the limits of T from 46 to 55 months. Group III contains cycles 12, 14 and 16 (limits from 56 to 65 months), and finally Group IV contains cycles 1, 5, 6 and 7, for which the period of ascension is longer than 70 months.

The main feature of this classification is that the secondary maxima and minima of spots within the cycles of each group are normally found at about the same intervals before and after the main maximum. This statement can be seen to be true from Figures 3a through 3d, on which the solid lines represent the mean value of the relative number R of the spots in each group, and the solid and stippled vectors represent the secondary maxima and minima in the cycles of each group.

If we should take into consideration not the relative numbers, but the area of the spots, we will obtain analogous results, which are shown in Figures 4a through 4c.

The same results will also be obtained if we establish a similar classification for the faculae (Figures 5a through 5c).

It should be noted that the shapes of the corresponding curves in Figures 3, 4 and 5 do not differ appreciably, even though the number of cycles analyzed for the areas of spots and faculae is smaller.

When the cycles are analyzed consecutively without regard to the period of ascension T, it is seen that the secondary maxima and minima are distributed at random. This is not the case when the cycles are analyzed in the order explained above. This fact poses the question of how proper are the reductions which normally obliterate these fluctuations and smooth out the sunspot curves. The solid line on Figure 6 represents the mean values of the differences of $R_{observed} - R_{reduced}$ for cycles 12, 14 and 16 of Group III; vectors above and below the abscissa axis represent the limiting positive and negative values of this difference for each cycle. It can be noticed that the limiting values of $R_{observed} - R_{reduced}$ for all the three cycles lie at the same period before and after the maximum, at a distance of one to two months. By reducing the observed values we systematically disregard these fluctuations and, consequently, we systematically obliterate the secondary maxima and minima at some moments.

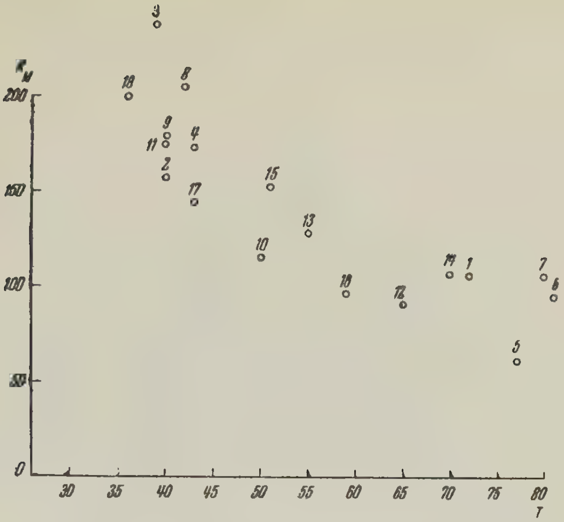


Fig. 1. Correlation between R_m and T .

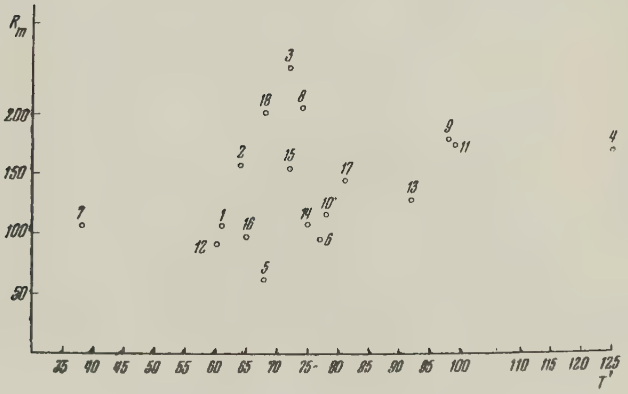


Fig. 2. Values of R_m and T'.

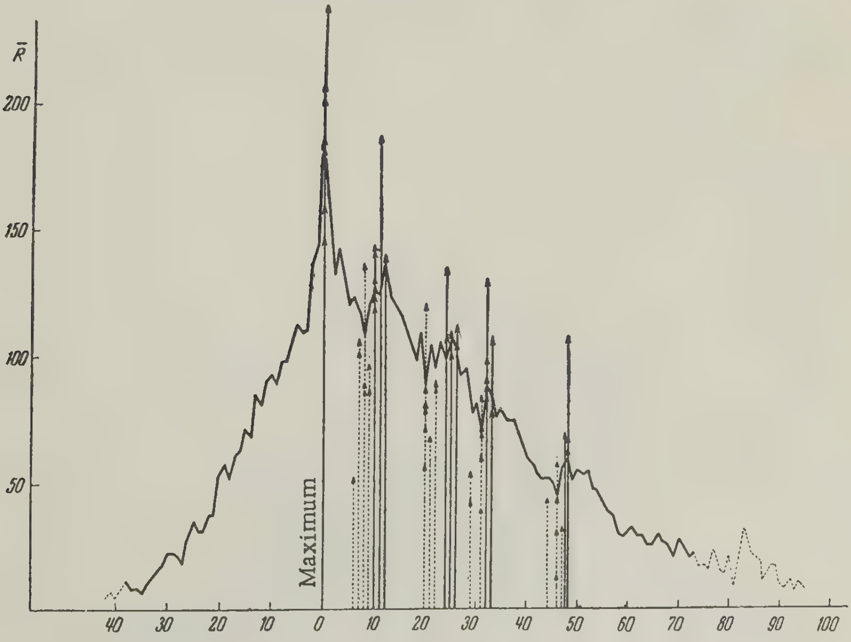


Fig. 3a. Group I, Cycles 2, 3, 4, 8, 9, 11, 17, 18.

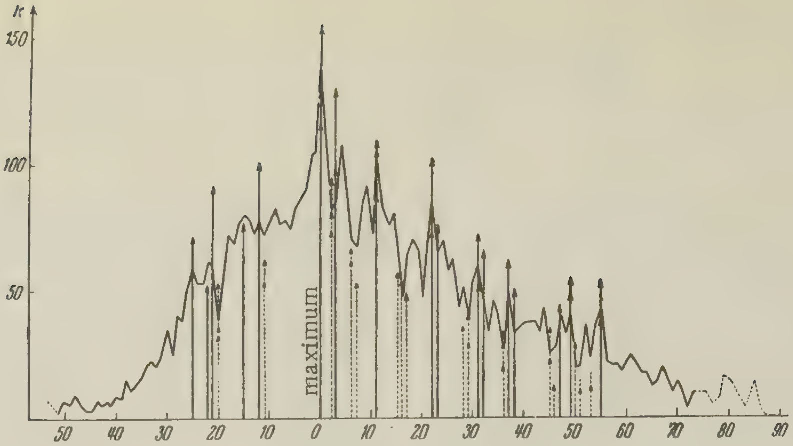


Fig. 3b. Group II, Cycles 10, 13, 15.

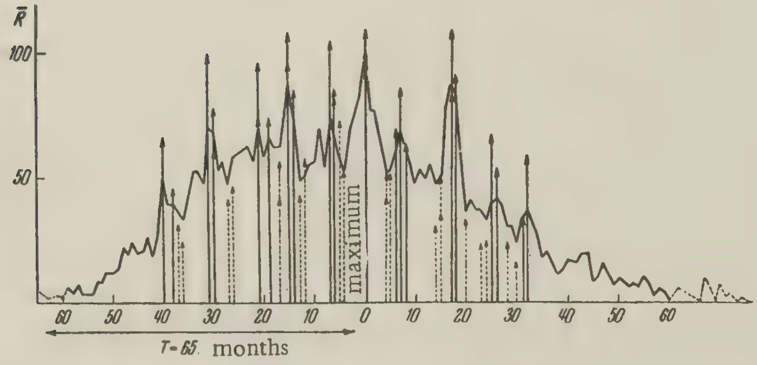


Fig. 3c. Group III, Cycles 12, 14, 16.

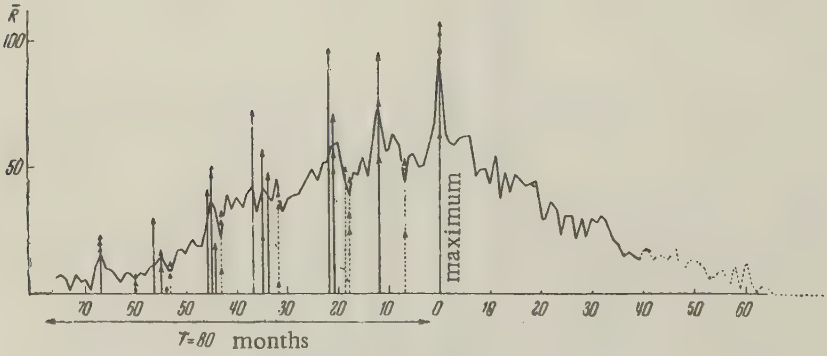


Fig. 3d. Group IV. Cycles 1, 5, 6, 7.

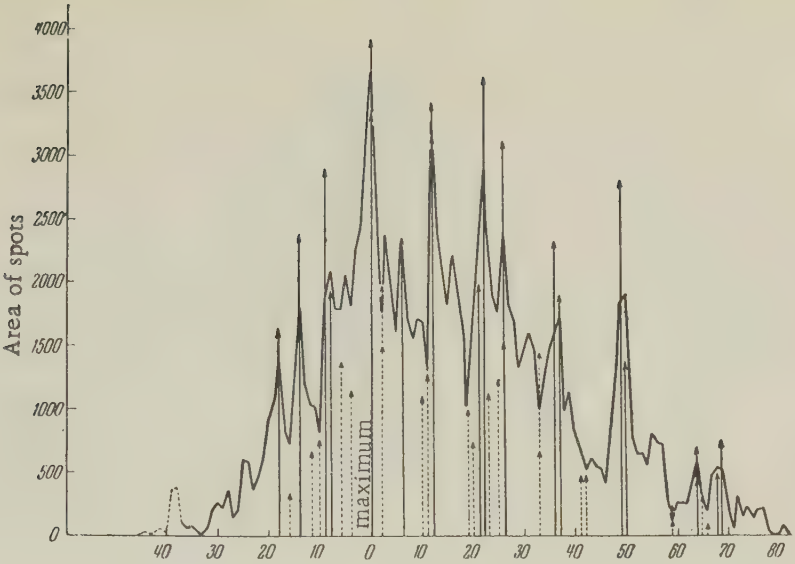


Fig. 4a. Group I, Cycles 17, 18.

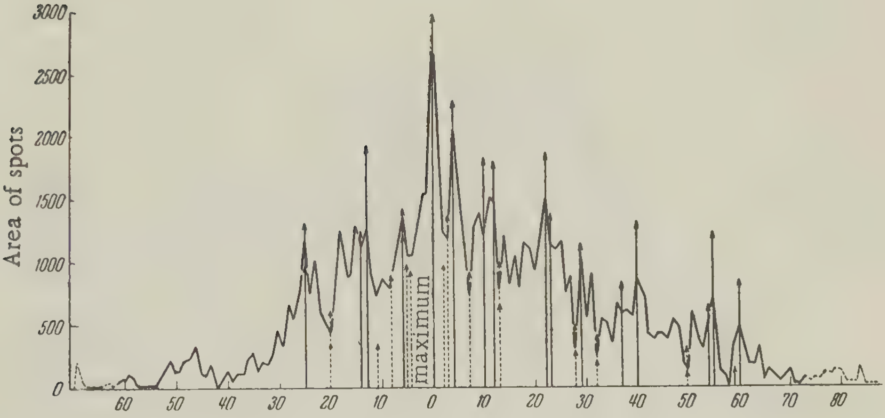


Fig. 4b. Group II, Cycles 13, 15.

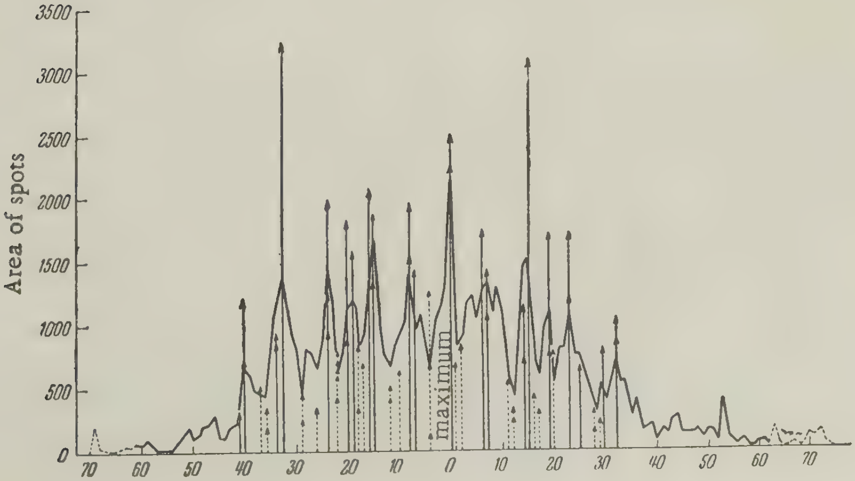


Fig. 4c. Group III, Cycles 12, 14, 16.

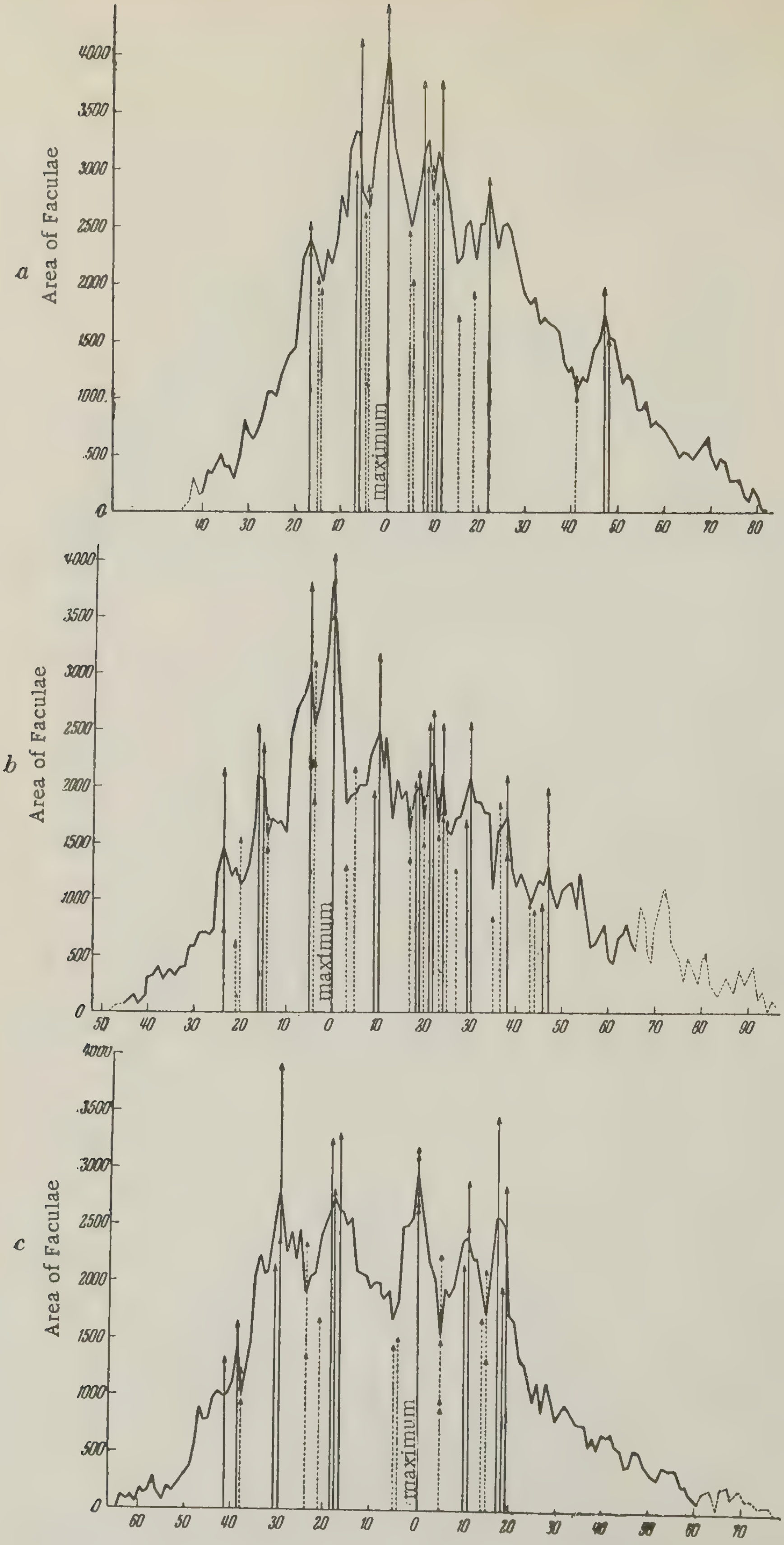


Fig. 5a. Group I, Cycles 17, 18; b) Group II, Cycles 13, 15; c) Group III, Cycles 12, 14, 16.

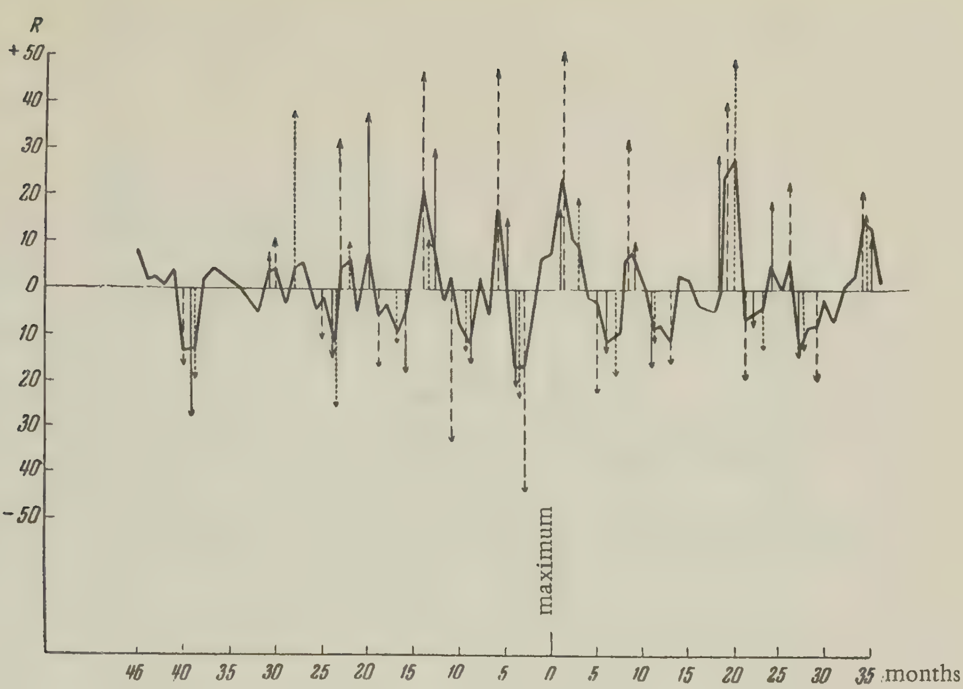


Fig. 6. Values of $R_{observed} - R_{reduced}$ for Group III.

It is then obvious that each solar cycle has a mechanism of its own, and that the fluctuations in the cycles of one group proceed with a more or less similar rhythm. This rhythm of fluctuations differs from one group to another, and for this reason the secondary maxima and minima appear to originate at random when we analyze the cycles consecutively or in some order which does not take into consideration the period of ascension T.

Special Features of the Groups

In connection with the above classification of the sunspot cycles into groups or types, we must make the following observations:

1) The ratio T/T' increases from the cycles of Group I through Groups II, III, IV (Figure 7). This is natural because the period of ascension T also increases from Group I through Group IV, while the period of decline T' decreases in the majority of cases.

2) Let S_1 represent the sum of the monthly mean values of the relative numbers R from the beginning of each cycle (a month after the minimum) through the month which includes the maximum, and let S_2 represent the corresponding sum from the month following the maximum through the month including the minimum. M. Waldmeier found from his work involving the reduced values that S_2 varies almost proportionally to R_m, while S_1 is almost independent of R_m (Figures 8a, 8b). It can be seen from Figures 9a and 9b that in our case S_2 is also nearly proportional to the observed R_m, while S_1 does not remain constant but slowly diminishes as R_m increases. If we substitute the period of ascension T for R_m, we will note that S_1 will slowly increase with T, while S_2 will vary in accordance with the parabolic law $S_2 = 1520 + 3.50(T - 75)^2$ (Figs. 10a and 10b).

3) Let N and S represent respectively the yearly mean area of spots in the northern and southern solar hemispheres, and let $(N/N + S)_G$ represent the mean value of the ratio of the sunspot area in the northern hemisphere to the total area of the sunspots for each group. Figures 11a and 11b show that this ratio behaves differently in Group I and in Groups II and III. Indeed, in Group I this ratio increases, exhibiting sinusoidal fluctuations, from the beginning to the end of the cycles, and assuming the value of $(N/N + S)_G \rightarrow 0.50$ near the maximum.

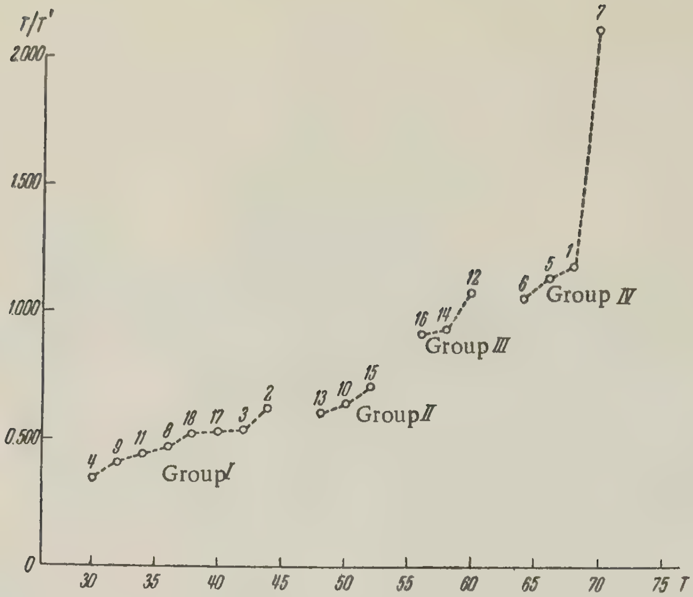


Fig. 7. Values of the ratio T/T'.

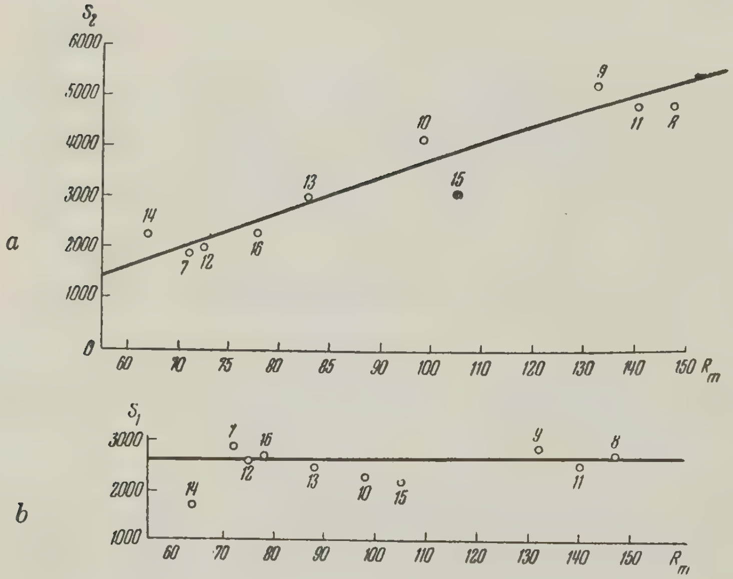


Fig. 8. Correlation between a) R_m and S_2; b) R_m and S_1, after M. Waldmeier.

On the other hand, in Groups II and III the ratio diminishes, again exhibiting sinusoidal fluctuations, and assuming the value of $(N/N + S)_G \rightarrow 0.45$ near the maximum. This indicates that at the beginning of the cycles in Group I, total area of sunspots in the northern hemisphere is, on the average, smaller than the total area of spots in the southern hemisphere. This asymmetry diminishes when the activity approaches its maximum, while near the maximum the total area of the spots in the two hemispheres becomes almost equal. As the cycles develop,

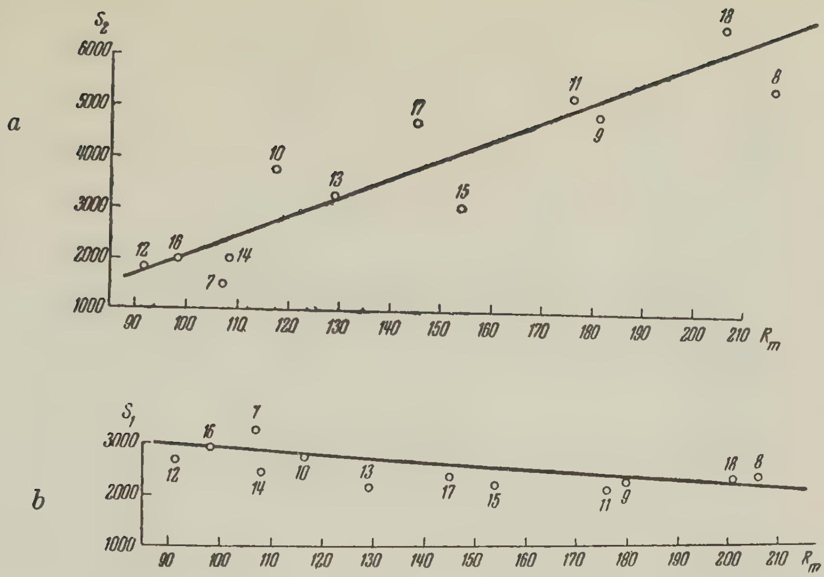


Fig. 9. Correlation between a) S_2 and R_m; b) S_1 and R_m.

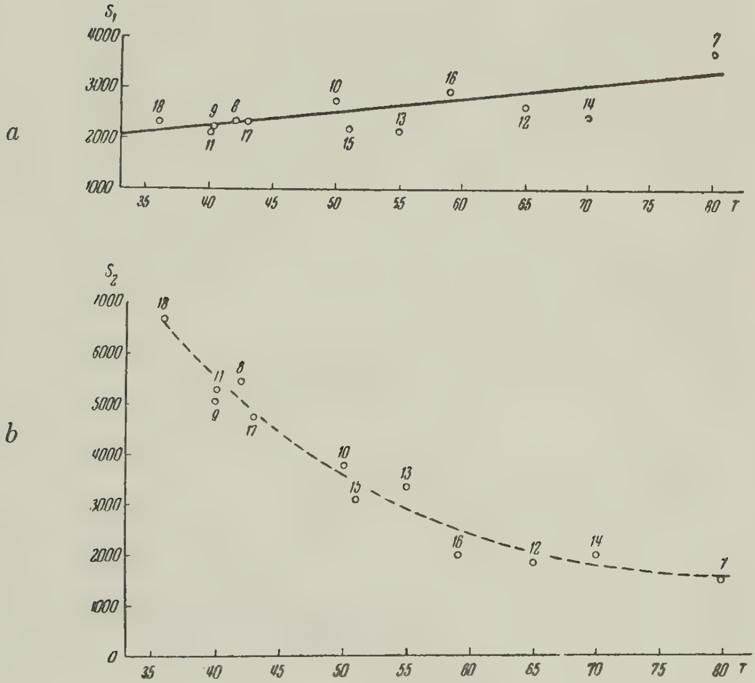


Fig. 10. Correlation between a) S_1 and T; b) S_2 and T.

the asymmetry alters so that the total area of spots in the northern hemisphere becomes larger than the total area in the southern hemisphere. The opposite case is true in Cycles II and III. This phenomenon is best shown in Groups I and III, while the cycles of Group II occupy an intermediate position.

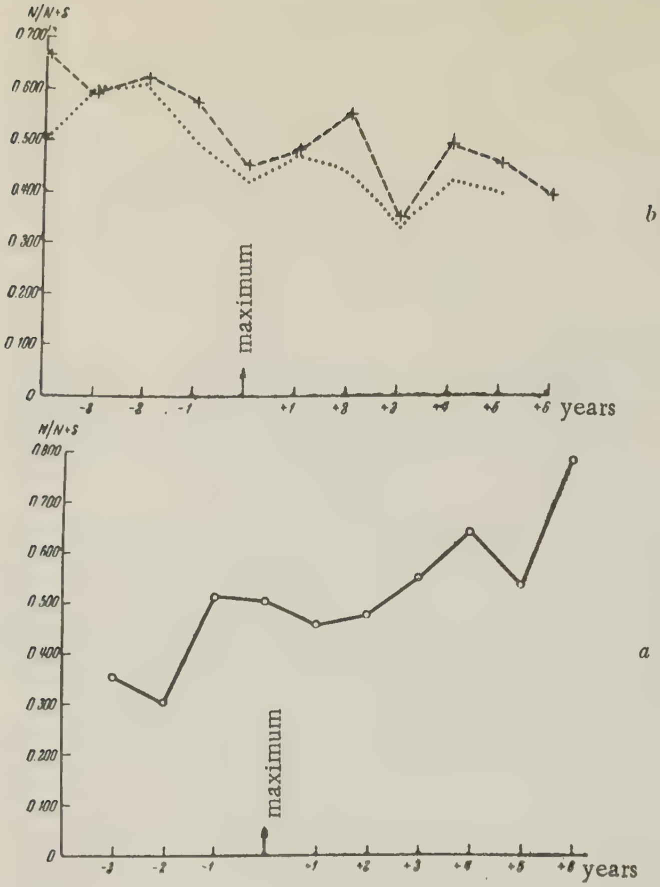


Fig. 11. Variation of N/N + S for the sunspots of a) Group I; b) Groups II and III.

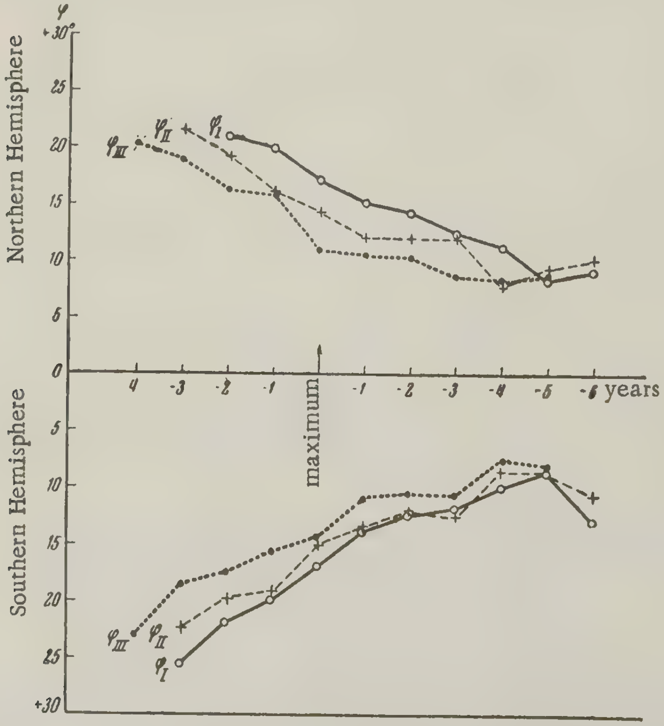


Fig. 12. Yearly variation in the mean latitude of the sunspots in each group.

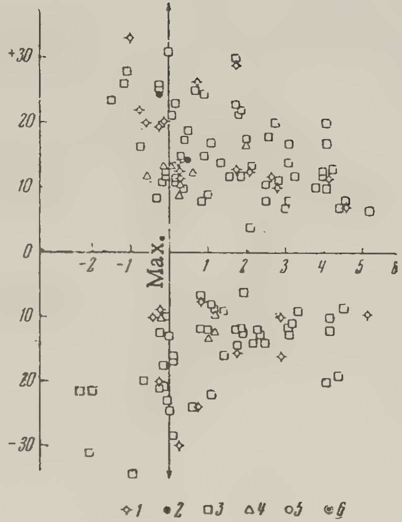


Fig. 13a.

Type I, cycles 17, 18
Northern Hemisphere 1-22 61%
Southern Hemisphere 1-14 39%
Northern Hemisphere 3-60 58%
Southern Hemisphere 3-43 42%

Weak disturbances (5 and 6) are disregarded.

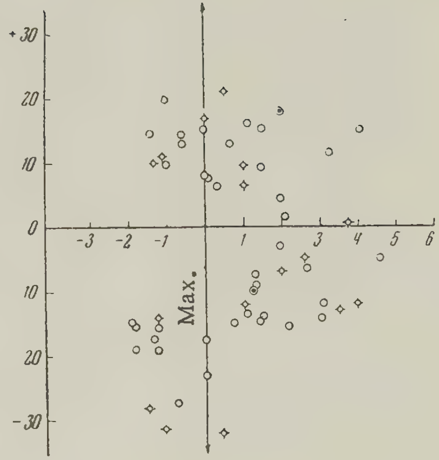


Fig. 13b.

Type II
Northern Hemisphere 1-7
Southern Hemisphere 1-9
Northern Hemisphere 5-18
Southern Hemisphere 5-22

Weak disturbances (3 and 4) are disregarded.

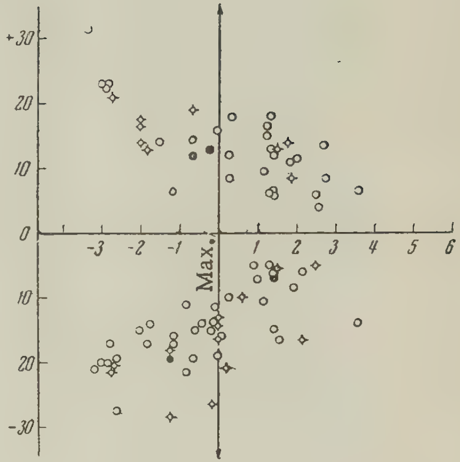


Fig. 13c.

Type III, cycles 12, 14, 16
Northern Hemisphere 1-10
Southern Hemisphere 1-14
Northern Hemisphere 5-30
Southern Hemisphere 5-41

Weak disturbances (4 and 3) were disregarded.

Legend: 1) Intense disturbance; 2) Intense disturbance, double; 3) Weak disturbance > 450; 4) Weak disturbance > 450, double; 5) Weak disturbances < 450; 6) Weak disturbance < 450, double.

4) The three groups behave differently with respect to the change of the yearly mean latitude of the spots before and after the maximum (Spoerer's law). Let φ_{I}, φ_{II} and φ_{III} represent the yearly mean latitude of the spots in the three groups. The solid lines in Figure 12 represent the values of φ_{I}, the change in the yearly mean latitude (northern and southern) for the spots in Group I. Lines shown by dashes give the values of φ_{II}, and the dotted lines give the values of φ_{III}.

Even though the differences between the changes φ_{I}, φ_{II}, and φ_{III} are not very large (2° - 3°), it is nevertheless obvious that the spots of Group I develop at higher latitudes than the spots of the other groups, and that their progress coincides only with the terminal part of the cycles, occurring four to six years after the maximum.

5) Figure 13 represents the latitudinal distribution of the spots with their area larger than 0.000, 500 of the solar hemisphere, which spots produced an intense or a weak ($3 > 450$, $5 < 450$) magnetic storm. It follows from the Figure that about 60% of magnetic storms in Group I are related to the spots of the northern solar hemisphere, and 40% to the spots of the southern hemisphere. The opposite relation is true for Groups II and III, in which the number of the intense and the weak magnetic storms produced by the spots of the northern hemisphere is smaller (41 to 44%).

Thus, with the help of our classification for the cycles of sunspots and faculae analyzed from the point of view of the period of ascension T, we can bring out some important features of the solar activity.

Academy of Athens, Greece

Received September 12, 1956

LITERATURE CITED

[1] Astr. Mit. der Sternwarte der Eidgen. Technischen Hochschule in Zurich, No. 133.

[2] Astr. Mit. No. 145, Zurich (1945).

[3] Royal Greenwich Observatory, London (1955).

ONCE MORE ON THE DISTANCES TO PLANETARY NEBULAE AND THE EVOLUTION OF THEIR NUCLEI*

I. S. Shklovskii

The objections raised by B. A. Vorontsov-Vel'iaminov to our system of distances of the planetary nebulae, and also to our point of view on their origin and evolution, are considered. It is shown that these objections carry no weight and are based on misunderstandings.

Numbers 2 and 3 of the Astronomical Journal for 1956 carry two of our papers [1, 2]. The first of them contains a revision of existing methods for determining the distances of individual planetary nebulae. In particular, in [1] it was emphasezed that the assumption made by B. A. Vorontsov-Vel'iaminov of a small luminosity dispersion for planetary nebulae implies that their optical thickness beyond the limit of the Lyman series is greater than unity. However, the process of expansion of the nebula that takes place inevitably causes the optical thickness beyond the Lyman series limit to become less than unity at some comparatively early stage of evolution. Thereafter the luminosity of the nebula will diminish rapidly. Thus the luminosity of "old" planetary nebulae must be many times smaller than that of "young" ones. If this is the case, a large dispersion of the luminosities of planetary nebulae is to be expected.

It is difficult to find any serious objection to these simple considerations, based on well-known laws of physics. In essence, B. A. Vorontsov-Vel'iaminov [3] recognizes this in making his criticism of our work. Nevertheless, he attempts to retard the continuous decrease of the optical thickness of the planetary nebula during its continuous expansion. Thus, for example, he tries to represent the expanding nebula as a shell with constant linear thickness, which leads to a slower decrease in luminosity proportional to r^{-2}, and in surface brightnes proportional to r^{-4}. It is easy to see, however, that the idea of an expanding nebula with constant linear thickness is not at all realistic. In view of all the structural complexity of the nebula, the process of its expansion will be quasi-homogeneous. The observed thicknesses of ring nebulae, for example, are only a few times smaller than their diameters.

The attempt of Vorontsov-Vel'iaminov to retard the rapid decrease of optical thickness beyond the Lyman limit by the hypothesis of the "sweeping up" of interstellar gas by the expanding nebula also does not withstand criticism. As grounds for this hypothesis, he points out that "at a moderate density of 10^{-22} g/cm^3 the interstellar gas will raise the mass of the shell by 10^{-2} $M_\odot$ when it expands to a radius of 50,000 A. U." Let us note, however, that one cannot call a density of 10^{-22} g/cm^3 for the interstellar gas "moderate" by any means. In fact, considering that the concentration of planetary nebulae within the galactic plane is substantially smaller than that of clouds of interstellar gas (the density of which $\sim 10^{-23}$ g/cm^3), we find that the overwhelming majority of planetary nebulae are surrounded by an interstellar medium having a density $< 10^{-25}$ g/cm^3. Thus this objection of Vorontsov-Vel'iaminov proves to be invalid. It must be admitted that the attempt of Vorontsov-Vel'iaminov to guarantee a continuous increase in the mass of the expanding nebula (which prevents a rapid decrease of its optical thickness) at the expense of corpuscular radiation from the nucleus of the planetary nebula, is also highly artificial. Unfortunately, Vorontsov-Vel'iaminov does not attempt a quantitative calculation, restricting himself only to theoretical considerations. Meanwhile a simple calculation shows that, in order to maintain the optical thickness of the expanding planetary nebula at a constant level, the output of corpuscular radiation from the nucleus must grow continuously in proportion to the cube of the time. This, obviously, is entirely inadmissible. Thus, the fact of the unrestricted expansion of planetary nebulae, the continuous decrease of their optical thickness beyond the Lyman series limit, and the rapid decrease of their luminosity, starting from a certain moment of their evolution, cannot be the object of much serious discussion. The whole question is, do the majority of observed planetary nebulae appear to be optically thin or optically thick ?

---

*An answer to B. A. Vorontsov-Vel'iaminov's criticism.

The method of distance determination proposed by us is useful for optically thin nebulae, which, in our opinion, substantiated in [1] and [2], constitute the majority of nebulae with known stellar magnitudes and angular diameters. This statement is disputed by Vorontsov-Vel'iaminov. He does not, for example, agree with the almost obvious condition that for the broad class of so-called "double-shelled" nebulae the optical thickness beyond the Lyman series limit must be less than unity. The arguments which he cites seem extremely unconvincing. Vorontsov-Vel'iaminov proceeds from the known fact that in the outer shells of double-shelled nebulae the luminescence depends principally on [OII], while the luminescence of hydrogen has an abrupt limit, practically coinciding with the radius of the inner shall. In the opinion of Vorontsov-Vel'iaminov this "is not consistent with the notion of incomplete absorption of ionizing radiation beyond the Lyman series limit" [3].

It is impossible to agree with this argument, however. Let us recall that the first ionization potential of oxygen (13.61 ev) is practically identical with the ionization potential of hydrogen (13.59 ev). Thus, in the second shell of the cloud, where the [OII] emission predominates, hydrogen must be completely ionized. The absence of observed hydrogen emission in the second shell is explained by the mode of its excitation (recombination), less effective than the mode of excitation of the oxygen ions (electron collision).

As soon as the hydrogen in the second, rather "faint", shell is found to be completely ionized, the double-shelled nebula has an optical thickness less than unity beyond the Lyman series limit. Vorontsov-Vel'iaminov recognizes the fact that, although the inner (bright) shell of the nebula may have an optical thickness less than unity beyond the Lyman limit, the outer (weak) one may, as it were, produce complete absorption. Again in this case he avoids quantitative calculations. Meanwhile, taking into account the relatively low surface of the second shell, the degree of ionization of hydrogen there will be very high and the optical thickness will be smaller than in the first shell. If the optical thickness of the second shell were greater than unity, it would "convert" a large part of the ultraviolet radiation of the central star. Consequently, its luminosity would be greater than that of the first shell, which is not observed.

Thus it must be recognized that the optical thickness of double-shelled nebulae beyond the Lyman series limit is less than unity. The objections advanced by Vorontsov-Vel'iaminov against this conclusion must be considered entirely unconvincing and based on misunderstandings.

To the number of optically thin nebulae belong also, in our opinion, nebulae with all kinds of "appendages" beyond the boundary of the main luminous mass, which is more or less compact (for example, the "Saturn" Nebula). Vorontsov-Vel'iaminov supposes, again purely speculatively, that the luminescence of these "appendages" is explained not by ultraviolet radiation from the nucleus beyond the Lyman series limit (and this would imply that the nebula was optically thin beyond the Lyman limit), but by "electromagnetic forces or other processes, as, for example, in the filamentary nebulae, which are devoid of a central star" [3].

Vorontsov-Vel'iaminov apparently has in mind processes of excitation by collisions on the high-temperature front of a shock wave (see, for example, [4]). It is well-known, however, that a characteristic of planetary nebulae is the low velocity of their expansion and internal motions. Here there can hardly be effects like those observed in filamentary nebulae (for example, in the remnants of supernovae).

There is no doubt that the very formation of "appendages", like all the complex morphology of planetary nebulae, is controlled to a considerable extent by magnetic fields. This was particularly stressed in our work (see [2]). But the luminescence of the "appendages", on the strength of the arguments cited above, must be explained by ultraviolet radiation from the nucleus of the nebula. Hence the optical thickness of such nebulae beyond the Lyman series limit must be less than unity. In our opinion, "ring" nebulae and irregular nebulae with low surface brightness should be counted among the planetary nebulae having an optical thickness less than unity beyond the Lyman limit. G. A. Gurzadian had before us substantiated this conclusion in detail [5]. Vorontsov-Vel'iaminov does not dispute the fact that this class of planetary nebulae belongs among those which are optically thin. Therefore we shall not again substantiate this conclusion here.

Analysis of existing data on planetary nebulae permits the conclusion that the majority (at least 75%) of nebulae with known stellar magnitude and angular diameters belong to one of the classes of double-shelled, ring, or faint semi-diffuse nebulae. This means that the majority of these nebulae are optically thin beyond the Lyman series limit. Of course, in individual cases there may be mistakes in assigning a particular observed nebula to the optically thin or the optically thick category. We understand this very well and emphasized it in [1]. Let us note, however, that the relative number of optically thin nebulae apparently was underestimated by us, as the criteria used were severe. In particular, this explains the fact, so troublesome to Vorontsov-Vel'iaminov, that objects with relatively low surface brightness fall in among the optically thick nebulae. The same cause (the severity of the criteria of selection) underlies the fact that we have not ranked the ring-shaped nebula BD + 30°3639 with the optically thin ones, since its surface brightness is very high. Such questionable cases are, however, very few, and the division of nebulae into two groups according to physically grounded criteria is carried out with sufficient certainty. Thus, the statement of Vorontsov-Vel'iaminov, concerning a supposed arbitrariness in assigning a particular nebula to the optically thick or optically thin category, lacks serious foundation.

Let us dwell now on the fundamental part of Vorontsov-Vel'iaminov's criticism, concerning our method of determining distances and on the connection of this method with others. With great regret we must note that Vorontsov-Vel'iaminov did not appraise our method in its essentials. The basic formula for the distance to an optically thin nebula has the form:

$$R \infty M^{2/5} / d'' I^{1/5},$$

where M is the mass of the nebula, d" is its angular diameter and I is the surface brightness Vorontsov-Vel'iaminov for some reason thinks that R is practically determined by d" as the reciprocal of the apparent angular diameter. In this connection Vorontsov-Vel'iaminov writes: "This means that Shklovskii is proposing a return to a very old method of distance determination, one that has, moreover, been rejected." It is easy to see, however, that so radical a statement by Vorontsov-Vel'iaminov does not correspond to the facts. According to his own data, cited in [3] the surface brightnesses of planetary nebulae vary within very broad limits, from 4^m to 19^m per square minute, i. e., by a factor of 10^6. Consequently, $I^{1/5}$ will vary from the faintest to the brightest nebulae by a factor of 16, i. e., by no small amounts.

Therefore, Vorontsov-Vel'iaminov's assertion that our method is a return to the old method of determining distances of planetary nebulae from their angular dimensions does not fit the facts. The surface brightness of an optically thin nebula, as well as its angular diameter, is a substantial factor in the determination of its distance. To the best of our knowledge no one before us has determined the distances of optically thin nebulae by such a method.

The extent to which Vorontsov-Vel'iaminov has not grasped our method of determining distances is evident from the fact that to him "it is not clear why Shklovskii gives the upper limit of the distances for them" (i. e., for optically thick nebulae. — I. Sh.) Meanwhile, and this was discussed in detail in [1], extrapolation of the relation of luminosity to linear radius r for an optically thin expanding nebula gives for small r improbably large values of the luminosity L_1. From this the conclusion may be drawn that in the early stages of evolution every planetary nebula must be optically thick. In that case its "true" luminosity L will be smaller than the extrapolated luminosity L_1. Hence, the distance of an optically thick planetary nebula must be smaller than that calculated from our basic formula, which is accurate for optically thin nebulae.

Furthermore, it surprises Vorontsov-Vel'iaminov that optically thick nebulae of small linear radius are comparatively infrequent (25%) among the known objects, although their luminosity and surface brightness are very large. One must not forget, however, that the evolutionary phase of an optically thick nebula proceeds comparatively rapidly, for approximately ten thousand years. The phase of an optically thin nebula lasts ten times as long.

The principal argument of Vorontsov-Vel'iaminov against our system of distances appears to be his statement about the enormous dispersion of the linear sizes of planetary nebulae. He has persistently adhered to this point of view for many years. This time Vorontsov-Vel'iaminov argues from the proper motions of planetary nebulae, which P. P. Parenago [6] obtained by averaging observations of various authors. Vorontsov-Vel'iaminov attempts to show that, for the different groups of nebulae distinguished by him, there is not a good correlation between the τ-components of the proper motions and the angular diameters. While the mean angular diameters for the three groups of nebulae distinguished by him are 14", 28", and 204", the corresponding values of $\bar{\pi}$ are $0''.0056$, $0''.0055$, and $0''.0090$. Since a change in the mean angular diameter of a nebula by a factor of 15 corresponds to a total change in $\bar{\pi}$ by a factor of only two, in his opinion the angular diameter does not appear to be a criterion of distance. Such a method of analyzing the distance criteria according to angular diameters would be supported if $\bar{\pi}$ for the group of nebulae with small angular diameters were equal to $0''.0006$. But one cannot expect the impossible: photographic astrometry cannot detect such small proper motions at the present time. Moreover, we take the liberty to make an extremely "heretical" statement: the real accuracy in determining proper motions of the nuclei of planetary nebulae, considering the circumstances of their observation, hardly exceeds $0''.003$ — $0''.005$.

After all, the planetary nebulae for which proper motions have been determined are very remote objects, found at a distance of 1-2 kps from us. A considerable portion, if not the majority, of comparison stars is closer to us. Therefore even a small motion of the centroid of the comparison stars leads to illusory proper motions for the nuclei studied. In this case we cannot expect great accuracy in the proper motions found for the nuclei.

What the mean proper motions of planetary nebulae are worth is apparent from the following example. According to [7], for the nebula NGC 7008, the following relative proper motions were found for the same comparison stars on five pairs of plates (Table 1a). From these individual results we "derive" the weighted mean $\bar{\mu}_x = +0''.0066 \pm 0''.0025$, $\bar{\mu}_y = -0''.0012 \pm 0''.0021$. Here is another example. According to [7] the following results are obtained for NGC 7662 on four pairs of plates (Table 1b):

TABLE 1a

μ_x	μ_y
$+0''.0173$	$+0''.0020$
-0.0132	-0.0088
$+0.0093$	-0.0129
$+0.0003$	$+0.0072$
$+0.0020$	-0.0102

TABLE 1b

μ_x	μ_y
$+0''.0070$	$+0''.0048$
-0.0126	-0.0006
$+0.0031$	$+0.0103$
-0.0045	-0.0004

But these nebulae belong to the number of relatively nearby objects. Vorontsov-Vel'iaminov absolutizes P. P. Parenago's absolute treatment of dissimilar series of observations, made by various authors with insufficient accuracy. Therefore we cannot be seriously concerned with his attempt to show that our distance scale is inconsistent with the proper motions. As was made clear above, the existing data on proper motions of planetary nebulae are entirely inadequate for the analysis of any scale of distances to planetary nebulae, in particular ours.

Only for a few of the nearest objects (for example, for the so-called "giant" planetary nebulae) are the proper motions real. In these cases they are found to be in good agreement with our scale.

Vorontsov-Vel'iaminov for some reason believes that "the observed unrestricted expansion of the nebulae must lead to enormous dispersion in their linear dimensions" [3]. This would be so, if in the process of their evolution the nebulae remained always optically thick. In fact at a rather early stage of its evolution the nebula has already become optically thin beyond the Lyman series limit. The majority of observed nebulae belong precisely to this last category of objects (see above). But, since the surface brightness of an expanding optically thin nebula falls quite rapidly, an increase in size of 4-5 times makes it practically unobservable. Of course, one may formally imagine the nebula expanding by hundreds of parsecs, whereupon its density will be several orders of magnitude smaller than the density of the interstellar medium. However, this hardly makes sense. That the dispersion in the linear sizes of observed nebulae is not great follows from the fact that their mean densities (under all assumptions concerning the distances) do not vary by more than a thousand times ($3 \cdot 10^2$ cm^{-3} < N_e < $3 \cdot 10^5$ cm^{-3}). Since $N_e \propto M/r^3$, for a small dispersion of masses the values of $\underline{r}$ for "large" nebulae (with small N_e) and for small, bright, dense nebulae differ only by an order of magnitude.

For lack of space we cannot dwell at any length here on certain other attempts by Vorontsov-Vel'iaminov to prove the correctness of his conclusion regarding a supposedly small luminosity dispersion of planetary nebulae and an enormous dispersion in their linear sizes. Insisting particularly on the correctness of his own interpretation of the empirical relation between surface brightness and angular sizes of the nebulae (the invalidity of this interpretation has been shown in [1]), Vorontsov-Vel'iaminov advances extremely strange arguments. The whole matter appears to be that "we (i. e., Vorontsov-Vel'iaminov – I. Sh.) are considering not an unlimited volume of space but a small volume of space surrounding the sun, in which the nebulae are within reach of discovery. Only for such a volume is it possible to determine the mean distance of the nebulae which Shklovskii himself is appraising" [3]. It is not comprehensible to us what relation this and the discussion following it have to the wholly specified problem of the smallness of the luminosity dispersion for planetary nebulae, as inferred from the diagram of "surface brightness vs. angular diameter". Entangling himself in attempts to counter our objection that this whole diagram is conditioned by observational selection, Vorontsov-Vel'iaminov, just a few paragraphs below, writes: "Thus, in the upper right corner of the diagram there are not and cannot be any nebulae in reality, and in the lower left corner they are unknown because they lie outside the sphere of observability." But here Vorontsov-Vel'iaminov in summarizing his own "objections", makes the same assertion we do.

Essentially, Vorontsov-Vel'iaminov has admitted that his diagram is an expression of observational selection. This being so, one must not draw from it far-reaching conclusions on the small luminosity dispersion of nebulae.

The dispersion of the absolute magnitudes of planetary nebulae may be obtained theoretically, proceeding from our assumptions about the evolution of these objects. Let us suppose that the luminosity of young, optically thick planetary nebulae is equal to L_0. At t_0 seconds after the formation of the nebula, the optical thickness beyond the limit of the Lyman series becomes equal to unity. Thereafter the luminosity of the nebula decreases considerably because of the continuing process of expansion with constant velocity, and at the instant of time $\underline{t}$ will be equal to $L = L_0 (t_0/t)^3$. We shall consider that the nebula may be observed up to the moment t_1, when its luminosity will have decreased to L_1. Then, as can be shown on the assumption that the formation of planetary nebulae proceeds at a constant rate, their luminosity distribution may be represented by the expressions:

$$dN(L) = \frac{\frac{1}{3} N\left(1 - \frac{t_0}{t_1}\right)}{\left[\frac{1}{L_1^{1/3}} - \frac{1}{L_0^{1/3}}\right]} L^{-4/3}\, dL \quad (\text{for} \quad L < L_0),$$

$$dN(L) = N \frac{t_0}{t_1} \delta(L - L_0)\, dL \quad (\text{for} \quad L = L_0),$$

where N is the total number of nebulae having luminosities in the interval L, L + dL, and $\delta\ (L - L_0)$ is the delta function.

The number of optically thick nebulae with constant luminosity (for which the distribution is represented by a delta function) constitutes a fraction t_0/t_1 of the total number of nebulae. On the basis of our work one may conclude that $t_0/t_1 \sim 1/5 - 1/10$.

The distribution of absolute magnitudes has the form:

$$N(M)\, dM \propto 10^{\frac{M-M_0}{2.512\cdot 3}}\, dM \cdot 10^{\frac{M-M_0}{7.54}}\, dM \quad (\text{for} \quad M > M_0),$$

$$N(M)\, dM \propto \delta(M - M_0)\, dM \quad (\text{for} \quad M = M_0).$$

The "accumulation" of nebulae with luminosity $L = L_0$ (or with absolute magnitude $M = M_0$ (see Figure 1) is a consequence of the schematization of the problem. In actuality there will be some (perhaps even a significant) dispersion of luminosities for the "young" optically thick nebulae being formed.

A dispersion in M_0 of even one or two stellar magnitudes liquidates the effect of "accumulation" of the nebulae in the neighborhood of $M = M_0$. Thereupon the distribution of nebulae takes the form plotted in Figure 2.

Consequently one may expect that over a broad range of absolute magnitudes N(M) will be approximately constant, and further, that N(M) will grow in proportion to the extent of the transition to "old" nebulae having relatively low surface brightness.

Recent observations by Koelbloed of planetary nebulae in the Small Magellanic Cloud [8] completely confirm this result of the theory. Figure 3a shows the observed absolute magnitude distribution for the brightest planetary nebulae in seven small areas of the Small Magellanic Cloud. Figure 3b shows the results of observations by

Lindsay [9] covering the whole Small Magellanic Cloud to $m_{pg} = 16.6$, while in [8] the limiting photographic magnitude was 18.1.

Thus the observations demonstrate the very great dispersion in absolute magnitude of the planetary nebulae, a fact which has a decisive significance for the problem under discussion.

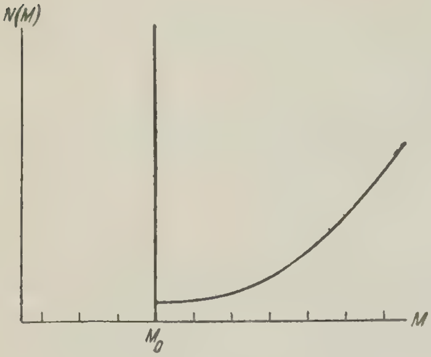


Fig. 1.

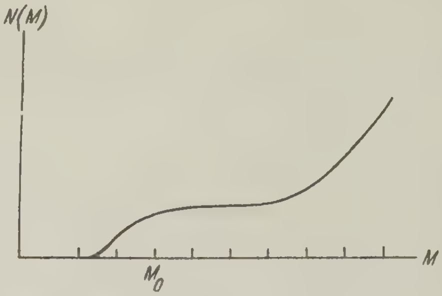


Fig. 2.

At the very end of his paper Vorontsov-Vel'iaminov criticizes our second work [2], based, to a considerable extent, on the new scale of distances to planetary nebulae contained in [1]. Since his objections to this scale are plainly unconvincing, it is hardly expedient to discuss here the sketchy and, it would seem to us, insufficiently thought out objections of Vorontsov-Vel'iaminov concerning our treatment of the problem of the origin and evolution of planetary nebulae. Moreover, in [2] we emphasized that our analysis is of the most preliminary character. It must be admitted, however, that two of our basic assumptions, which Vorontsov-Vel'iaminov for some reason passed over in silence in his criticism, are sufficiently well-founded. The first is that the nuclei of certain planetary nebulae (e. g., NGC 7293) are superdense white dwarfs having very high temperatures. This important conclusion which we have drawn from the new distance scale is obtained directly from the estimated masses of the nebulae. If these nebulae (e. g., NGC 7293) were further away, their masses would appear to be excessively large. Therefore it seems reliably established that, as a result of the formation of the planetary nebula, at least in a few cases the nucleus is transformed into a white dwarf. Allowing for the fact that planetary nebulae represent a positively related group of objects (according to origin and evolution) the connection between white dwarfs and planetary nebulae hardly seems accidental. Another matter that is difficult for us at present is a well-defined description of the basic features of the decidedly rapid process of evolution of the nuclei of planetary nebulae. We have shown in the most preliminary manner the possible paths of this evolutionary process. This, let us note, is an entirely secondary part of our work, perfectly debatable, and the criticism of Vorontsov-Vel'iaminov will perhaps be useful in the future. Another very important conclusion drawn in [2] concerns the "ancestors" of the planetary nebulae. As it seems to us, it was successfully demonstrated that planetary nebulae must be formed from certain peculiar giant stars with low surface temperature. The detachment of an envelope from such a star with practically zero velocity depends upon internal mechanisms of stellar evolution, and may be closely connected with the problem of the structure and evolution of certain types of red giants. This, our most important conclusion, for some reason remained outside the scope of Vorontsov-Vel'iaminov's critical paper. Most recently, after publication of our papers, other works have appeared in which an idea is developing on the relation of cool giants and white dwarfs that is rather similar to our own [10]. However, as far as we know, no one as yet has considered with sufficient clarity the problem of the genetic relation of certain types of cool giants, planetary nebulae, and white dwarfs.

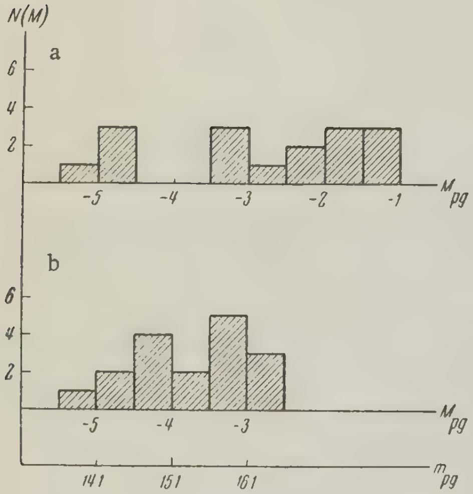


Fig. 3.

Thus, as was shown above, Vorontsov-Vel'iaminov's criticism of our work on planetary nebulae is based on a series of misunderstandings and, in its turn, does not withstand criticism. It is important to emphasize that in the past two decades astrophysics has been enriched by new ideas and concepts that permit advances to be made on the difficult problem of the origin and evolution of planetary nebulae, a problem until very recently left in an entirely unsatisfactory state. Under the circumstances it is hardly worthwhile to cling to old ideas.

P. K. Shternberg State Astronomical Institute

Received December 24, 1956

LITERATURE CITED

[1] I. S. Shklovskii, Astron. J. (USSR) 33, 2, 222 (1956).

[2] I. S. Shklovskii, Astron. J. (USSR) 33, 3, 315 (1956).

[3] B. A. Vorontsov-Vel'iaminov, Astron. J. (USSR) 33, 6, 809 (1956).

[4] S. B. Pikel'ner, Izv. Krymsk. Astrofiz. Obs. 12, 93 (1954).

[5] G. A. Gurzadian, The Dynamics of Planetary Nebulae [in Russian] (Erevan, 1954).

[6] P. P. Parenago, Astron. J. (USSR), 23, 69 (1946).

[7] J. Anderson, Lick. Obs. Bull. 17, 21, 462 (1934).

[8] D. Koelbloed, Observatory 76, 191 (1956).

[9] E. M. Lindsay, Monthly Notices, Roy. Astron. Soc. 115, 248 (1955).

[10] A. Deutsch, Astron. J. 61, 174 (1956).

RESULTS OF PHOTOGRAPHIC OBSERVATIONS OF MARS AT THE KHARKOV ASTRONOMICAL OBSERVATORY DURING 1954

I. K. Koval'

On the basis of results of observations of Mars, made in four spectral regions on the 20 cm refractor of the Kharkov Astronomical Observatory from June 1 to September 10, 1954, the author comes to the following conclusions:

1. For the reduction of observations of Mars the formula of V. V. Sobolev and the formulae of V. G. Fesenkov and E. Schönberg may be used with equal success.

2. The slope of the brightness distribution curves in the spectral region 560-660 mμ decreases with increasing meridian altitude of the sun.

3. Near the date of opposition, the color of the "maria" differs little from the color of the "continent." The "maria" at low southern latitudes change their color with variation of the meridian altitude of the sun.

4. The reflection of light from the surface of the "maria" of Mars does not follow Lambert's law.

5. The atmosphere above the Martian "maria" is more transparent than that over the "continent".

6. The south polar cap is reddish in comparison to a white screen.

At present a series of works is being published, in which the authors refer to the results of an investigation of Mars obtained at the Kharkov Astronomical Observatory (KhAO) in recent years. In connection with these we consider it expedient to supplement data published earlier [1-4] by presenting the results of studies of Mars carried out in 1955 on materials of the 1954 observations.

Photographic observations of Mars were made from June 1 to September 10, 1954 on the 8-inch refractor of the KhAO. The transmission maxima for the filters used were the following: red — 640 mμ, yellow — 580 mμ, green — 520 mμ, blue — 460 mμ. The plate scale was 14" per mm. The photographs were standardized on the star Arcturus.

During each night of the observations the coefficient of transparency of the earth's atmosphere was determined by the "long method".

The calibration of the plates was made with diffuse sunlight. All plates were developed under standard conditions. For two images on each plate, measurements were made on the MF-2 microphotometer. The polar regions, the regions of the "maria" clearly visible on the photographs of Mars, and the "continent" along the equator of intensity and at the borders of the polar regions, were measured. Also measured were the extra-focal images of the star and the scale calibration marks.

It is important to note that our observations cover the beginning and middle of the Martian spring for the southern hemisphere. During the time of observations the meridian altitude of the sun for points on the surface of Mars varied by 18°.

As a result of the reduction of the observations made in the vicinity of opposition (June), curves of brightness distribution along the equator of intensity were obtained (values of the brightness factor ρ. Table 1).

TABLE 1

i	640 mμ	580 mμ	520 mμ	460 mμ
0°	0.247	0.204	0.169	0.127
10	0.245	0.203	0.169	0.127
15	0.242	0.199	0.167	0.126
20	0.237	0.194	0.164	0.125
25	0.227	0.188	0.161	0.123
30	0.217	0.181	0.156	0.121
35	0.207	0.174	0.151	0.119
40	0.195	0.164	0.145	0.117
45	0.183	0.156	0.140	0.115
50	0.168	0.145	0.134	0.112
55	0.155	0.136	0.126	0.110
60	0.139	0.124	0.117	0.107

On the basis of the data of this table the fundamental parameters which characterize the optical properties of the atmosphere and surface of Mars were determined with the aid of the formula of V. V. Sobolev [5] in the form

$$r = \chi(\gamma) r_0 + \Delta r + [r - r + r(q)], \tag{1}$$

where r is the brightness coefficient; r_0 is the brightness coefficient for a spherical indicatrix; $\chi(\gamma) r_0$ is a term depending on the first-order scattering; $\chi(\gamma)$ is the indicatrix of scattering in the earth's atmosphere, found by V. V. Sobolev; Δr is a term that takes into account scattering of orders higher than the first; and $[r - \bar{r} + r(q)]$ is a term depending on the presence of ground reflecting light according to the law:

$$f(i, \varepsilon) = \cos^q i \cos \varepsilon.$$

Here

$$r(q) = \frac{2Ae^{-\tau \sec i} \cos^{q-1} i - 1) R(\varepsilon)}{4 + (3 - \chi_1)(1 - A)\tau}.$$

The determination of the albedo A, the smoothness factor q for the surface, and also the optical thickness τ of the atmosphere oi Mars, was performed by the method of least squares. The data are listed in Table 2.

TABLE 2

λ (mμ)	640	580	520	460
A	0.240	0.186	0.152	0.104
τ	0.011	0.021	0.036	0.064
q	0.98	0.98	0.99	1.04

With the values of τ obtained, the term Δr, at a point having an angle of incidence of 60° for the blue filter, is equal to 0.004, i. e., it does not exceed the error of the method of photographic photometry itself. This allows us to discard the rather cumbersome Δr term in Formula (1).

Therefore it is no wonder that the theoretical curves of brightness distribution plotted on the basis of the formulae of V. V. Sobolev [5], V. G. Fesenkov [6] and E. Schönberg [7], equally well fit the observed brightness distribution with values of A and τ chosen from Table 2.

We also obtained curves of brightness distribution for each month of the observations (Table 3).

TABLE 3

i	June 7	July 9	Aug. 15	Sept. 10	i	June 7	July 9	Aug. 15	Sept. 10
	Red filter					Green filter			
0°	0.243	0.247	0.248	0.246	0°	0.168	0.169	0.166	0.163
10	0.242	0.245	0.247	0.245	10	0.168	0.169	0.166	0.163
15	0.239	0.242	0.244	0.242	15	0.166	0.168	0,165	0.162
20	0.234	0.236	0.239	0.239	20	0.164	0.166	0.163	0.160
25	0.225	0.227	0.232	0.234	25	0.160	0.162	0.159	0.156
30	0.216	0.218	0.223	0.226	30	0.155	0.157	0.155	0.152
35	0.206	0.207	0.213	0.218	35	0,149	0,151	0.149	0.147
40	0.194	0.195	0.202	0.208	40	0.143	0.144	0.143	0.142
45	0.182	0.184	0.200	0.197	45	0.136	0.137	0.137	0.137
50	0.165	0.169	0.176	0.184	50	0.130	0.130	0.130	0.132
55	0.151	0·158	0.164	0.172	55	0.123	0.122	0.123	0,126
60	0.134	0.143	0.149	0.158	60	0.114	0.113	0.115	0.119
	Yellow filter					Blue filter			
0°	0.205	0.205	0.207	0.206	0°	0.126	0.125	0.123	0.121
10	0.204	0.204	0.206	0.205	10	0.126	0.125	0.123	0.121
15	0.201	0.202	0.203	0.202	15	0.125	0.124	0.122	0.120
20	0.196	0·197	0.199	0.198	20	0.124	0.123	0.121	0.119
25	0.188	0.190	0.194	0.194	25	0.122	0.122	0.120	0.118
30	0.180	0.183	0.187	0.189	30	0.120	0.120	0.119	0.117
35	0.172	0.175	0.180	0.184	35	0.118	0.118	0.117	0.115
40	0.164	0.166	0.172	0.177	40	0.115	0.115	0.115	0.113
45	0.155	0.157	0.165	0.170	45	0.113	0.113	0.113	0.111
50	0.145	0.148	0.155	0.162	50	0.110	0.110	0.111	0.109
55	0.134	0.137	0.145	0.153	60	0.102	0.103	0.104	0.103
60	0.123	0.127	0.135	0.142					

From this table it follows that the slope of the curves for the red and yellow filters decreases with increasing meridian altitude of the sun. There is almost none of this in the green and blue regions of the spectrum.

Evidently the degree of dustiness of the atmosphere of Mars grows with the approach of summer in the Martian southern hemisphere.

We investigated 18 regions situated in the "maria" of Mars. All of these regions, according to the variation of their spectral reflecting powers, are markedly distinguished into the following three groups.

1. Regions situated with the zone of lattitude $-21° < \varphi < +30°$. This group is characterized by a marked decrease of albedo in the red and yellow regions of the spectrum, and almost unvarying albedo in the green and blue, with increasing meridian altitude of the sun (Table 4). In color, the "maria" of this group becomes less red in comparison to the color of the "continent".

2. The group of "maria" lying in the zone of latitude $-60° < \varphi < -25°$. The "maria" of ~~this~~ group are distinguished by a marked decrease of albedo over the whole spectrum with increasing meridian altitude of the sun. The color of these "maria" did not vary (Table 5).

3. The regions of the "maria" with latitudes $\varphi > +30°$. These zones during the whole period of observation did not vary either in albedo or in color (Table 6).

At the start of our observations (near opposition) the color of the "maria" under study in the southern hemisphere was close to the color of the "continent" (Tables 4-6).

TABLE 4

"Mare" Syrtis Major; $\varphi = -8°$; $\lambda = 288°$

Date	640 mμ	580 mμ	520 mμ	460 mμ
		$A_{\text{м}}$		
June 10	0.208	0.182	0.153	0.081
July 16	0.193	0.165	0.148	0.084
August 18	0.169	0.150	0.140	0.083
		$A_c/A_{\text{м}}$		
June 10	1.155	1.010	0.990	1.143
July 16	1.243	1.093	1.250	1.106
August 18	1.420	1.200	1.085	1.115

TABLE 5

Date	640 mμ	580 mμ	520 mμ	460 mμ
		$A_{\text{м}}$		
June 16	0.228	0.162	0.143	0.104
July 16	0.216	0.148	0.131	0.087
August 19	0.191	0.134	0.117	0.082
September 10	0.183	0.131	0.115	0.080
		$A_c/A_{\text{м}}$		
June 16	1.054	1.114	1.065	0.980
July 16	1.110	1.220	1.160	1.190
August 19	1.255	1.340	1.306	1.260
September 10	1.311	1.413	1.310	1.300

TABLE 6

Date	640 mμ	580 mμ	520 mμ	460 mμ
		$A_{\text{м}}$		
June 16	0.221	0.173	0.129	0.092
August 13	0.226	0.161	0.143	0.097
September 16	0.218	0.168	0.130	0.095
		$A_c/A_{\text{м}}$		
June 16	1.137	1.042	1.162	1.127
August 13	1.062	1.097	1.063	1.061
September 16	1.081	1.107	1.119	1.100

This is in agreement with the spectroscopic observations of Mars conducted in 1954 by N. A. Kozyrev [8], who, as is known, obtained a parallel run of spectral curves for the "continent" and "maria", and also with photographic observations made in the same year by Kuiper on color film [9].

In regard to the smoothness of the surface of Mars the following should be stated. It is possible to form an opinion about the degree of smoothness of a planet's surface only on the basis of a study of the brightness distribution over its disk (if we are concerned with the smoothness of the whole surface of the planet, which is more or less homogeneous in its optical properties) or on the basis of a study of the variation in brightness of surface details with variation in the angle of incidence of the light (if we are concerned with the smoothness of certain regions located on the surface of the planet).

The smoothness of the surface of Mars as a whole may be judged on the basis of a study of the brightness distribution over its disk in the red and infrared if we neglect the insignificant effect of the atmosphere of Mars on the observed brightness distribution curves obtained at these wavelengths.

At the present time the overwhelming majority of astronomers who are studying Mars (it is understood that we are speaking of astronomers who are studying Mars directly by methods suitable for quantitative interpretation) come to the conclusion that the surface of the "continent" of Mars is similar in its reflecting properties to surfaces that reflect light according to Lambert's law. The smoothness factor for the "continent" of Mars is close to unity. More complex is the task of determining the smoothness factor for the "maria" of Mars. It must be said that one may conjecture about the smoothness of the surface of the Martian "maria" on the basis of a study of the brightness variation of a specific region of a "mare" as it departs from the central meridian (observing Mars in the course of 4-5 hours on one night or at the same time in the course of 4-5 nights). The smoothness factor $\underline{q}$ which we found in this way for the "maria" of Mars proved to be the same for all "maria", remained unchanged during the whole period of observation, and equals 0.62 [3].

A rough idea of the surface roughness of the Martian "maria" in general may be obtained by studying in the red and infrared the averaged brightness distribution curves for the "maria" visible on the disk under various angles of incidence of light. It is impossible, however, to draw any conclusion whatever regarding the smoothness of individual regions of the Martian surface on the basis of visual observations. On a visual scale it is not possible to estimate the degree of decrease in brightness of a "mare" with its displacement toward the edge of the disk, much less ascribe to this brightness variation a specific law (for example, Lambert's law). One may talk about the change in contrast of the "continent" and "mare" brightnesses observed visually in red light as the "mare" is displaced from the center to the edge of the disk of Mars.

In the following table are listed the spectral values of the smoothness factor $\underline{q}$ for the "mare" Syrtis Major, as well as values of the optical thickness of the atmosphere lying above this "mare".

TABLE 7

λ (mμ) Date	640	580	520	460
q				
June 10	0.60	0.61	0.57	0.62
July 6	0.56	0.60	0.57	0.60
August 18	0.61	0.65	0.61	0.56
τ				
June 10	0.007	0.019	0.031	0.044
July 6	0.008	0.016	0.030	0.043
August 18	0.008	0.017	0.027	0.042

As is evident from this table, the optical thickness of the atmosphere above a Martian "mare" has, according to our observations, somewhat lower values than that over the "continent" (this conclusion also refers to a remaining investigation of the "maria" of Mars being undertaken).

It is important to note that the smoothness factor for the "mare" Syrtis Major, which we found on the basis of the brightness distribution in infrared light (840 mμ) on the 6/7 September 1956 (Figure 1) turned out to be + 0.50 i. e., it remained the same as in 1954.

Probably the surface roughness of the "mare" Syrtis Major does not have seasonal fluctuations.

As for the polar regions of Mars, the observations which we made above the south polar zone lead to the data of Table 8.

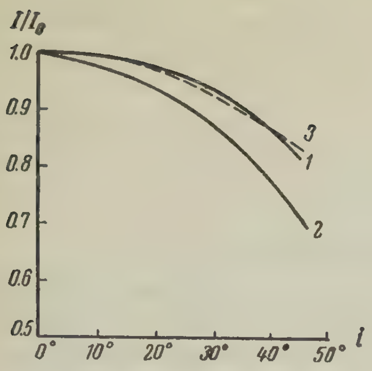


Fig. 1.

In conclusion let us list the results for A and q obtained from E. Schönberg's formula for "continent" and "mare" of Mars, using the τ_λ found by N. A. Kozyrev. The results are given in Table 9. Here τ_s and τ_a are the scattering and absorption components of the optical thickness of the Martian atmosphere.

From this table it follows that in the case considered the whole surface of Mars must have an azure tint. The smoothness factor in this case too maintains a constant value along the spectrum.

From all that has been said one may draw the following conclusions:

1. For reducing the observations of Mars one may use with equal success the formula of V. V. Sobolev, or the formulas of V. G. Fesenkov and E. Schönberg.

TABLE 8

λ (mμ)	640	580	520	460
$A_{\text{м}}$	0.402	0.373	0.306	0.182
$A_c/A_{\text{м}}$	0.597	0.500	0.500	0.571
τ	0.011	0,019	0.051	0.080
q	0.67	0.72	0.93	1.11

TABLE 9

λ (mμ)	640	580	520	460
		Continent		
τ_λ	0.100	0.380	0.980	1.380
τ_s	0.020	0.080	0.255	0.428
τ_a	0.080	0.300	0.725	0.952
A	0.248	0.243	0.333	0.342
q	0.70	0.74	0.72	0.78
		Mare		
A	0.124	0.122	0 166	0.171
q	0.55	0.54	0.56	0.59

2. The decrease in the slope of the brightness distribution curves in the red and yellow portions of the spectrum with increasing meridian altitude of the sun leads one to think that, as summer approaches, the power of the Martian atmosphere to scatter red light grows. Evidently, with the approach of summer in the southern hemisphere the amount of dust in the atmosphere of Mars increases.

3. The absence of a difference in color of the majority of "maria" and the "continent" of Mars, disclosed by several observations at the time of opposition of Mars in 1954, is confirmed in this work and explained, apparently by the presence of a winter-spring season of the year in the southern hemisphere of Mars at the moment of observation.

4. The color of some of the "maria" of Mars depends on meridian altitude of the sun. As the meridian altitude of the sun increases the "maria" darken (the albedo falls over the whole spectrum) and some of them

become less red in comparison to the "continent". One may suppose that with the approach of summer the "maria" are moistened through the thawing of the polar cap and that in some of them there are regions occupied by Martian vegetation. N. P. Barabashov came to such a conclusion on the basis of his many observations at KhAO.

5. The smoothness factor for the "continent" and "maria" has different values, each the same over the whole spectrum, which evidently points to a difference in surface structure for the "maria" and the "continent" of Mars. Reflection of light from the surface of the "maria" does not follow Lambert's law.

6. The atmosphere above the "maria" of Mars is more transparent than that above the "continent".

7. The south polar cap is reddish in comparison to a white screen. This result confirms those conclusions which were obtained at KhAO in previous years. The maximum value of the smoothness factor for the surface of the south polar cap in the blue region of the spectrum agrees well with the assumption of N. P. Barabashov regarding the fact that the polar caps of Mars represent a thin layer of frost or ice above which at a low altitude float sparse cloudlike formations.

Kharkov Astronomical Observatory

Received November 5, 1956

LITERATURE CITED

[1] N. P. Barabashov, Astro. Tsirkuliar KhAO, No. 11 (1953).

[2] I. K. Koval', Astr. Tsirkuliar AN SSSR No. 159 (1955).

[3] I. K. Koval', Astr. Tsirkuliar KhAO, No. 15 (1956).

[4] N. P. Barabashov and I. K. Koval', Astron. J. (USSR) 33, No. 6 (1956).

[5] V. V. Sobolev, Usp. Astr. Nauk 6 (1954).

[6] V. G. Fesenkov, Astron. J. (USSR) 24, 257 (1944).

[7] E. Schönberg, Theoretische Photometrie, Hdb. d. Astroph., T. 11/1.

[8] N. A. Kozyrev, Izv. Krymsk. Astrofiz. Obs. 4 (1956).

[9] G. P. Kuiper, The Atmospheres of the Earth and Planets [Russian translation] (IL, 1951), p. 382.

ON THE PROBLEM OF DETERMINING THE MASS DISTRIBUTION OF METEORS

E. I. Fialko

A simple method for a more precise determination of the mass distribution of meteors by means of radar observations is discussed in this paper.

INTRODUCTION

A knowledge of the mass distribution of meteors is necessary for the solution of several very important problems: the determination of the density of meteoric matter in interstellar space, the investigation of the influence of meteoric ionization on the state of the ionosphere (in particular, the verification of the meteoric hypothesis of the production of the sporadic E-layer at intermediate latitudes) and on radio communications, etc.

As is known, the mass distribution of meteors is usually taken to be a power law [1, 2]:

$$f_m(m) = \frac{b}{m^S}, \tag{1}$$

where m is the meteor mass, b and S are constants.

The value of S can be found from visual observations; however, considerable corrections have to be introduced because of the subjective nature of the visual observations [2].

Moreover, visual observations can only yield the distribution for comparatively bright meteors, while it is more important to determine the distribution in the region of small masses. This is all the more essential since, from observations on the Zodiacal Light, it is found that the number of small particles in interplanetary space is considerably greater than the number given by (1) with the values of S obtained from visual observations.

By means of radar observations we can eliminate this contradiction and not only determine the value of the exponent S more accurately, but also obtain a more accurate expression for the meteor-mass distribution which, apparently, is described by a more complicated function than Equation (1).

Kaiser [2, 3], using the experimental data of McKinley [4], found that $S \approx 2$ for sporadic meteors, while for the Quadrantid stream $S \approx 1.7$ (previously, Lovell [5] found that for the Quadrantids $S \approx 1$).

In these experiments [4], the power of the transmitter was varied perodically and also two radar stations were used with different power outputs and with slightly different wavelengths.

The S can be obtained from the frequencies observed at two stations which use significantly different wavelengths, but in this case the results need correcting, since the dependence of the observed meteor frequency on the wavelength is not accurately known.

The determination of S by a comparison of the meteor frequencies obtained by radar stations with different transmitter power outputs requires the use of two (or more) stations.

The determination of S by means of one station with variable transmitter power-output is associated with technical difficulties if the power output has to be varied over a wide range (for the study of a wide range of meteor masses).

Because of these considerations, it is desirable to discuss a simple and workable method for determining S by the use of one radar station.

§ 1. Basic Equations

Kaiser [2] has shown that the ratio of the meteor frequencies registered by two sets of radar equipment, which have antenna systems with identical directivity patterns, can be found from the approximate equation:

$$\frac{N_1}{N_2}=\left[\frac{P_1\lambda_1^3}{\varepsilon_1}\,\frac{\varepsilon_2}{P_2\lambda_2^3}\right]^{(S-1)/2}, \tag{2}$$

where N is the average number of meteors registered by the radar per unit time (for example, per hour), P is the peak power radiated by the transmitter, λ is the wavelength, ϵ is the power of the threshold signal, and the subscripts 1 and 2 refer to the first and second radar stations, respectively.

Therefore, if we have two (or more) radar stations with different parameters, we can find the value of S from a comparison of the frequencies. However, the same result can be obtained if we have one radar station with two receivers (having different sensitivities) or one receiver with an adjustable signal threshold:

$$\frac{N_1}{N_2}=\left[\frac{\varepsilon_2}{\varepsilon_1}\right]^{(S-1)/2}, \tag{3}$$

which yields

$$S=1+2\,\frac{\log\frac{N_1}{N_2}}{\log\frac{\varepsilon_2}{\varepsilon_1}}. \tag{4}$$

In Equations (2)-(4), N_1 (or N_2) is the number of meteors which can produce an echo exceeding the threshold signal.

However, some of the meteors will not be registered by the radar despite the fact that the electron density α may be sufficient to give a reflected signal greater than the noise level when the meteor trail is irradiated by radio waves.

This can happen in the case of meteors of short duration, during pulse operation of the equipment, if the pulse repetition interval is longer than the time in which the trail exists.

If the pulse repetition rate is low, a considerable number of short-duration meteors will not be observed for this reason.

The following question naturally arises: do we introduce an appreciable error if, in finding S from Equation (4), we use the frequencies of observed meteors ?

Bullough has shown that [6]

$$N'=N\zeta,$$

where N is the number of meteors which can produce an echo of amplitude greater than the threshold, N' is the frequency of observed meteors, and ζ is a quantity given by the expression

$$\zeta=P^{1-S}+\frac{1}{\ln P(s-1)}\left[1-P^{1-S}\{1+(S-1)\ln P\}\right];\quad \text{where}\quad P=e^{T_i/T_{tr}},$$

here T_i is the pulse repetition interval, T_{tr} is the duration of the reflection (which would be observed with a continuous irradiation of the trail) calculated for a level which is 1/e as large as the maximum amplitude:

$$T_{\mathrm{tr}} = \frac{\lambda^2}{16\pi^2 D};$$

where D is the coefficient of diffusion at an altitude corresponding to the position of the reflecting section of the meteor trail and λ is the wavelength.

Therefore, the quantity ζ in the first approximation is independent of the magnitude of the threshold signal. Hence

$$\frac{N_1}{N_2} \approx \frac{N_1'}{N_2'}$$

and

$$S \approx 1 + 2\,\frac{\log \frac{N_1'}{N_2'}}{\log \frac{\varepsilon_2}{\varepsilon_1}}. \tag{5}$$

In order to find out the mass distribution of meteors, it is necessary to determine what range of mass values corresponds to the observed value of S. For this, it is necessary to know the smallest mass of an observable meteor.

§ 2. Minimum Mass of a Detectable Meteor

The detection of a meteor, produced by a body of minimum mass, will occur under the most favorable conditions for detection, namely:

1. The linear electron density is at a maximum at the point of intersection of the meteor trail with the normal from the observing point on to the meteor trail (Fig. 1).

2. The slant range (from the observing point to the normally-reflecting section of the trail) and the inclination of the trail (the zenith angle of the radiant) are the most favorable (see below).

3. The geocentric velocity of the meteor is at a maximum, etc. As is known [7], the maximum electron density $\alpha_{\max}$ in a meteor trail is

$$\alpha_{\max} = \frac{4}{9}\, m\beta \cos\chi\, [\mu H]^{-1}, \tag{6}$$

where m is the initial mass of the meteor, β is the probability that one atom evaporated from the meteor produces one free electron (as the result of a collision with the atoms of the air), χ is the zenith angle of the meteor radiant, μ is the mass of one atom of the meteor, and H is the height of the homogeneous atmosphere.

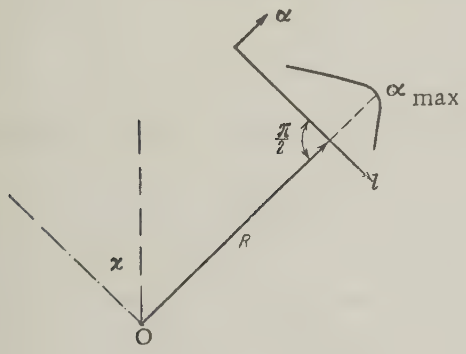


Fig. 1. Favorable conditions for the detection of a meteor. O) Location of the detector; R) slant range; χ) zenith angle of the meteor radiant; α) linear electron density; l) is measured along the meteor trail.

The received echo power for the case $\alpha < 2.4 \cdot 10^{12}$ electrons/cm will be [2, 3]

$$\varepsilon_{\mathrm{in}} = \frac{P_i (G\eta)^2 S^4(\theta)\, \lambda^3}{16\pi^2 R^3}\, \alpha^2 \gamma^2, \tag{7}$$

where P_i is the peak power radiated by the transmitter, G is the directive gain of the antenna, $S(\theta)$ is the directional-response pattern of the antenna, λ is the wavelength, R is the slant range, η is a constant depending on the orientation of electric-field vector of the incident wave with respect to the meteor trail, $1 < \eta < 2$ ($\eta = 1$ for parallel polarization, $\eta \approx 2$ for perpendicular polarization).*

$$\gamma = \frac{e^2}{m_e c^2},$$

---

*Recent data indicate that η can also be greater than 2 [8].

where e is the electronic charge, m_e the electron mass, and c the velocity of light.

From Equations (6) and (7) it follows that the minimum mass of a detectable meteor is given by

$$m_{\min} = \left[\frac{\varepsilon}{P_i G^2 \lambda^3} \frac{1}{\beta^2_{\max}} \frac{1}{\eta^2_{\max}} \left(\frac{R^3 H^2}{S^4(\theta) \cos^2 \chi} \right)_{\min} \frac{16\pi^2 \mu^2}{\gamma^2} \right]^{1/2}, \tag{8}$$

where ϵ is the threshold-signal power for the receiver.

In calculating the quantity

$$\left(\frac{R^3 H^2}{S^2(\theta) \cos^2 \chi} \right)_{\min}$$

we have to take into account that

$$R = \frac{h}{\sin \chi}$$

$$H = H(h),$$

where h is the altitude of the normally-reflecting section of the meteor trail.

The coefficient β, determining the probability for ionization, is a function of the meteor velocity:

$$\beta \sim v^n.$$

The magnitude of the exponent n has not yet been established. According to the data of some authors $n < 1$, but in the opinion of other authors the value of n is considerably greater than unity and $n = 5.6$ [9].

Thus the minimum mass of a detectable meteor is determined by the detector parameters (λ, ϵ, P_i, G, $S(\theta)$), as well as by the meteor parameters (v, μ).

From (8) it follows that, in order to obtain the mass distribution $P_m(m)$ in the region of small meteor masses, it is necessary to work at comparatively long wavelengths λ, to use a high-directional antenna and a high-power transmitter.*

§ 3. Some Methods for Varying ϵ

Several methods can be used for the reception of signals reflected from meteor trails under real conditions. We note the following two methods:

1. The method of multichannel reception — we use several receivers (or one multichannel receiver) whose signal thresholds are not equal.

2. The method of single-channel reception — one receiver is used and the signal threshold is varied periodically.

The variation of the signal threshold can be accomplished by means of either a noise generator or a voltage divider (attenuator).

Without carrying out a comparison of the two methods here, let us note that the version to be preferred is the single-channel receiver with a noise generator that is periodically switched on and off (for example, the generator is switched on for 1 minute, then it is disconnected for another minute); one cathode-ray oscillograph with an intensity-modulated linear sweep (along a diameter) is used; the recording is by means of a photographic film which moves with a uniform low speed (for example 1 cm/sec) in front of the indicator screen, at right angles to the time base sweep.

When the noise generator is switched on, the film is appropriately marked.

---

*A further discussion of $m_{\min}$ will be published elsewhere.

To determine N'_1 and N'_2, we count separately the number of meteor echoes registered on the film when the noise generator is switched on and when it is switched off.

Since the interference level changes (for example, in the course of a day), it is necessary to measure the power ϵ_1 continually and to alter ϵ_2 in an appropriate manner.

Having determined N'_1 and N'_2 and the ratio of the threshold signals ϵ_1/ϵ_2, we find S from Equation (5).

S. M. Kirov
Tomsk Polytechnical Institute

Received December 1, 1956

LITERATURE CITED

[1] B. Iu. Levin, Physical Theory of Meteors and Meteoric Matter in the Solar System [in Russian] (Izd. AN SSSR, 1956).

[2] T. R. Kaiser, Monthly Notices, Roy. Astron. Soc. 114, 39 (1954).

[3] T. R. Kaiser, Monthly Notices, Roy. Astron. Soc. 114, 52 (1954).

[4] D. W. McKinley, Can. J. Phys. 29, 5 (1951).

[5] A. Lovell, Uspekhi Fiz. Nauk 11, 1 (1950).

[6] K. Bullough, Jodrell Bank Annals, ser. I, I, 68 (1954).

[7] T. R. Kaiser, Advances in Phys. 2, 8 (1953).

[8] E. R. Billam and J. C. Browne, Meteors,(T. R. Kaiser, editor) (1955), p. 73.

[9] G. S. Hawkins, Astrophys. J. 124, 311 (1956).

APPLICATION OF SILVER SULFIDE PHOTOELECTRIC CELLS WITH A BARRIER LAYER TO ABSOLUTE SPECTROPHOTOMETRY AND PYROMETRY

I. MEASURING APPARATUS USING A SILVER SULFIDE PHOTOCELL

G. F. Sitnik

Absolute measurements have certain specific features which must be taken into account in the design of such measurements.

An apparatus using a silver sulfide photocell is described and consists of a combination of a compensation short circuit and a potentiometric device. The characteristics of the apparatus, an account of tests on it, and a description of the method whereby the components are chosen, as well as the precautions which have to be taken to improve the accuracy, are described.

The apparatus may be used to compare luminous fluxes, varying by a factor of 400 or more, with the same sensitivity. The deviation from direct proportionality between the incident luminous flux and the photocurrent is 0.5-0.7% on the average. The apparatus has been in use in absolute measurements during some seven years.

For a number of years now, absolute spectrophotometric measurements of the brightness of laboratory sources and of the sun have been carried out at the Kuchin Astrophysical Observatory. These measurements have been carried out by comparison with a standard source in the form of a perfect black body specially constructed for this purpose [1-3]. All the spectrophotometric and pyrometric measurements were carried out using photocells. The present paper consists of a description of an experiment on absolute measurements using a silver sulfide photocell with a barrier layer (FESS-U) which was developed and prepared at the Institute of Physics of the Academy of Sciences of the Ukrainian SSR [4-7].

The real advantage of photocells with barrier layers is that they produce a photocurrent without any additional power supplies. Furthermore, silver sulfide photocells have a high integral sensitivity (up to 700 μa/lu) and sufficiently high sensitivity in the red and near infrared. The latter property is particularly important in pyrometric and spectrophotometric studies.

There are two main disadvantages as far as the FESS-U are concerned and they must be taken into account in absolute measurements.

1. Deviations from direct proportionality between the photocurrent and the luminous flux. A. A. Il'ina [8] has shown that some of the FESS photocells exhibit a deviation from direct proportionality of 5-6%, or even more. Better examples of these photocells gave smaller deviations, up to 0.5-1.5%, when illumination was varied by a factor of 10. An increase in the latter factor leads to a bigger deviation from the direct proportionality between luminous flux and the photocurrent; the photocurrent increases faster than the luminous flux. V. E. Kosenko and E. G. Miseliuk [6] have shown the presence of deviations from direct proportionality (of the order of 1.5%) under the short circuit regime. The latter authors have carried out some measurements on the FESS-U in the range 100-200 luxes. An increase in the resistance of the external circuit of the photocell leads to even bigger deviations from the direct proportionality between the photocurrent and the luminous flux.

2. The temperature dependence of the photocurrent. Within the temperature range − 20 to + 30° C the photocurrent decreases by 0.5-2% when the temperature increases by 1° C [6-8]. Our own measurements on the FESS also show that the photocurrent depends appreciably on temperature. Such a strong dependence of the sensitivity of the FESS on temperature is particularly objectionable in absolute measurements which necessarily take considerable amount of time to carry out.

Nevertheless the successful application of FESS in scientific studies, and in particular in spectrophotometric studies [8-10], has established the possibility of the application of these photocells to absolute measurements. Our problem was to choose a measuring circuit which together with suitable precautions would fully neutralize these defects of the FESS-U. In addition, the measurements had to be limited to a small range of illumination which fortunately is the case in spectrophotometric practice. The work on the application of the FESS to absolute measurements was begun in 1948.

In the present paper a measuring circuit used in conjunction with the FESS-U and consistent with all the main requirements associated with absolute measurements [1-3] is described. A brief account of this apparatus was given earlier [11]. Here a more detailed description is given of the apparatus, tests on it, and the choice of the components, as well as of practical experience gained during the use of the setup.

It is well known that photocells with barrier layers show a deviation from the linear dependence between the incident luminous flux and the photocurrent which is due to a change in the resistance of the barrier layer under the action of the incident flux and for other reasons [12, 13]. Such a departure from linearity sets in the sooner, the greater is the load resistance in the external circuit. This external resistance is shunted by the low variable resistance of the barrier layer. As a result of this, the primary photocurrent is divided into two parts: one part I', flows through the barrier layer while the other, I, through external resistance r. It is easy to show [12] that the current I, which flows through the external circuit, is given by the following relation:

$$I = \frac{k\Phi}{1 + \frac{r_2 + r}{r_1}}, \tag{1}$$

where r_1 is the resistance of the barrier layer, r_2 is the resistance of the remaining parts of the photocell (i. e., of the semiconductor, metallic electrodes and contacts), $r_1 + r_2$ is the total resistance of the photocell between the external terminals, Φ is the incident luminous flux, and k is the integral sensitivity of the photocell.

When r = 0 we obtain the short circuit current for which we have the following relation:

$$I_0 = \frac{k\Phi}{1 + \frac{r_2}{r_1}}. \tag{2}$$

In the majority of cases the resistance r_2 is negligibly small compared with r_1. If this is so, the short circuit current I_0 depends linearly on Φ, i. e., $I_0 \approx k\Phi$.

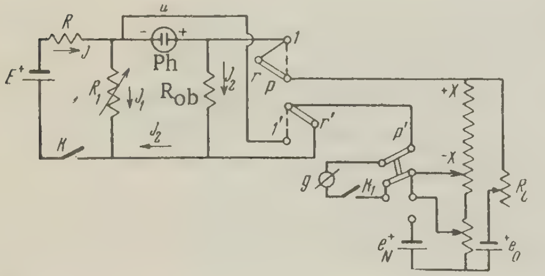
There is an internal escape path for the primary photocurrent through the variable resistance of the barrier layer. For this reason there is a departure from an exact linearity between the external current I, and sometimes even the short circuit current I_0, on the one hand, and the incident flux on the other.

The effect of the variable resistance of the barrier layer r_1 on the external current I, and consequently on the departure from linearity of the photocell scale, could be neutralized by preventing the escape of the current through the barrier layer i. e., by establishing a zero potential difference across the barrier layer itself. However, this condition cannot be achieved exactly. It can be achieved approximately by having the potential difference across the outer terminals equal to zero throughout the measurements. This principle was first suggested by Compball [14] and Wood [13]. Later, electronic circuits with feedback [15] were based on this compensation method. Such a method insures that the photocell works, in effect, on short circuit even when the external load is quite considerable. In this way the photocell characteristic is improved, and, above all, the direct proportionality between the photocurrent and the luminous flux is preserved to a sufficient degree even for large external loads.

The compensation method, i. e., the method using zero potential difference between the output terminals of the photocell is particularly good in combination with a potentiometer. Such a combination may be used to measure, with the same intensity, luminous fluxes which change quite considerably. This is very important in absolute measurements. Furthermore, among the contemporary measuring instruments the potentiometer insures maximum accuracy.

The circuit diagram of a measuring device incorporating a silver sulfide photocell Ph which was employed by us in setting up a model of a high temperature black body is given below [1-3]. The potential difference applied to the photocell is tapped off the variable resistance R_1. This potential difference is opposite in sign to the electromotive force of the photocell (photoelectric emf). The resitsance R_1 was in the form of a 10,000 ohm five dial decade resistance box of type MSR-47. This resistance box in conjunction with a suitably chosen resistance R and the emf of the battery E gives a continuous and accurate compensation of the photoelectric emf, u. An external low resistance R_{ob} is in series with the photocell having a resistance $r_{ph} = r_1 + r_2$. After a zero potential difference has been established across the output terminals of the photocell, short circuit current flows across the resistance R_{ob}. By measuring the change in the potential difference across this resistance we obtain a quantity which is proportional to the short circuit current and, consequently, to the luminous flux incident on the photocell.

The low resistance potentiometer PN-2,class one, was used to measure the potential differences; the potentiometer was developed in 1948 (cf. [16] and [17]). A potential difference may be applied to the terminals X of the potentiometer through the key P either from the resistance R_{ob} or from the photocell Ph. A highly sensitive galvanometer of type Tsernike C was used to indicate the zero point (absence of potential difference). The measuring scale was at a distance of 6 meters from the galvanometer. This galvanometer, in accordance with the principles given in [16] and [18], is very suitable for work in conjunction with the low resistance potentiometer shown in the circuit diagram below. A change of one unit in the last decade of the low resistance potentiometer ($\Delta e = 1 \cdot 10^{-6}$ v) gives a change of one unit in the deflection on the scale.



Measuring a circuit incorporating a silver sulfide photocell.

In practice the measurement of the photocurrent using this curcuit was carried out as follows. First of all it is necessary to compensate the photoelectric emf across the terminals of the photocells and due to the luminous flux by means of the external potential difference. In order to achieve this the key P (see figure) is set in position 1-1'. All the decades of the potentiometer are then set to zero. In other words, the galvanometer g and the "zero resistance" of the potentiometer are connected across the photocell. By means of the key K_1 which is included in the potentiometer we apply the photoelectric emf to the terminals of the galvanometer. For large luminous fluxes the same precautions must be taken as in the case of the primary compensation of the working current of the potentiometer [18]. Having achieved an approximate compensation of the photoelectric emf, and keeping it connected in, we change the resistance R_1 until the crosswire of the galvanometer falls on the zero of the scale. This means that the potential difference across the illuminated photocell is now zero. In other words, the short circuit condition has been satisfied and the short circuit current is flowing through the resistance R_{ob}. The key P is then set into the position r-r'. The potential difference across the resistance R_{ob} is then measured using the potentiometer and is proportional to the photocurrent and the luminous flux incident on the photocell. With some training, one measurement is carried out in two to three minutes. This slow rate at which the measurements can be made is a disadvantage of the method. However, in absolute measurements the accuracy and reproducibility which are obtained by this method are more important than the rapidity with which the measurements can be taken.

Let us now consider the problem of the best choice of the components in the above measuring circuit. For this purpose it is necessary to obtain certain relationships which characterize the above circuit. It is easy to see from the above figure that in the general case, i. e., in the absence of the compensation of the photoelectric emf, we have for the circuit including the photocell, R_1 and R_{ob},

$$u - r_{ph} I'_2 = I'_2 R_{ob} - I'_1 R_1, \tag{3}$$

$$I = I'_1 + I'_2. \tag{4}$$

Here $\underline{u}$ is the emf of the photocell when illuminated, r_{ph} is the resistance of the FESS for given illumination and load, I'_1 and I'_2 are the currents through R_1 and the photocell respectively, and I is the total external current of the source E given by the relation

$$I = \frac{E}{R + \frac{R_1 (R_{ob} + r_{ph})}{R_1 + R_{ob} + r_{ph}}} \tag{5}$$

By solving Equations (3) and (4) we have

$$I'_2 = I \frac{R_1}{R_1 + R_{ob} + r_{ph}} + \frac{u}{R_1 + R_{ob} + r_{ph}}, \tag{6}$$

i. e., the current flowing through the load consists, in general, of two components, namely, the photocurrent corresponding to this resistance and the additional current due to the external source.

When the compensation has been achieved i. e., when the potential difference across the photocell is zero, instead of Equation (3) we have

$$u - r_{ph} I_2 = I_2 \cdot R_{ob} - I_1 R_1 = 0. \tag{7}$$

By analogy with (6) we then have

$$I_2 = \frac{u}{r_{ph}} = \frac{R_1}{R_{ob}} I_1 = I \frac{R_1}{R_1 + R_{ob}}. \tag{8}$$

Here I_2 denotes the short circuit current and I_1 the current through the resistance R_1 on short circuit. With accurate compensation of the photoelectric emf across the photocell the potential difference across R_1 is equal to the potential difference across R_{ob}, i. e.,

$$R_1 I_1 = R_{ob} I_2.$$

The circuit elements must be chosen so that accurate and continuous compensation of the photoelectric emf across the terminals of the photocell can be achieved. Let us consider the choice of the circuit elements in an actual case as illustrated in the above figure. The voltage sensitivity of the measuring scheme was, as was shown above, 10^{-6} v/mm of the scale. A continuous compensation of the photoelectric emf is achieved if a change $\Delta R_1 = 1$ ohm in R_1 gives a change in the potential difference across the photocell corresponding to tenths of a millimeter on the scale. An error in the compensation of $\Delta R_1 = 1$ ohm will lead to a change in the short circuit current of ΔI_2. From (8) we have the following expression for the latter quantity:

$$\Delta I_2 = I \frac{\Delta R_1}{R_1 + R_{ob}} \frac{R_{ob}}{R_1 + R_{ob}}. \tag{9}$$

Putting $R_{ob} = 1 \cdot 10^5$ ohm, $\Delta R_1 = 1$ ohm, and R_1 between 1 and 10,000 ohm into Equation (9), we find the following limiting values for ΔI_2: $1.00 \cdot 10^{-5} I$ and $8.26 \cdot 10^{-6} I$. These limiting values of ΔI_2 do not differ from each other by very much. We shall take $\Delta I_2 = 1 \cdot 10^{-5} I$. In that case $\Delta\omega$, the error in the determination of the change in the potential difference across R_{ob}, will be $\Delta\omega = R_{ob} \Delta I_2 = R_{ob} 1 \cdot 10^{-5} I$. Suppose this error is $\Delta\omega' = 4 \cdot 10^{-6}$ v. Then $I = 4 \cdot 10^{-6}$ amp and $\Delta I_2 = 4 \cdot 10^{-11}$ amp. In accordance with Equation (7), the error in the compensation of the photoelectric emf across the photocell will be

$$\Delta u = (r_{ph} \Delta I_2) = 1.2 \cdot 10^{-7} \mathrm{V} \tag{10}$$

(r_{ph} was taken equal to the mean value of the resistance of the FESS-U2 which was 3,000 ohm). With such a change in the potential difference, continuity and accuracy of compensation is insured.

Using the value $I = 4 \cdot 10^{-6}$ amp, one can choose the values of E and R in accordance with Equation (5). This value of the current is obtained, for example, for $E = 2.25$ v and $R = 3 \cdot 10^5\,\Omega$. With such values of the parameters of the apparatus, small luminous fluxes can be reliably measured. However, in the case of large currents, it may turn out that the introduction of even the whole resistance R_1 will not fully compensate the emf of the photocell. In other words, in this case the necessary current may, in accordance with Equation (8), turn out to be greater than $I_2 = 0.091 \cdot 4 \cdot 10^{-6}$ amp $= 3.6 \cdot 10^{-7}$ amp. If this is so then the compensation of the photoelectric emf may be achieved by increasing the emf of the battery E, or by decreasing the resistance R. The most convenient method of changing the current I is by changing R, which in the present experiments was in the form of a resistance box whose total resistance could be increased up to 1 MΩ in steps of $10^5\,\Omega$. For the increased value of the current I, the continuity and accuracy of compensation of the photoelectric emf is ensured by the introduction of the fifth decade of the resistance box R_1 in steps of 0.1 Ω. Correspondingly, the potentiometer may, with equal accuracy, be used to measure photoelectric currents corresponding to large and small luminous fluxes. To measure very weak currents we use a resistance $R = 6 \cdot 10^5\,\Omega$.

Let us now consider the choice of the resistance R_{ob}. At first sight it may seem that the resistance R_{ob} could conveniently be quite large. This however, is incorrect. To begin with, if the external resistance is too large the short circuit regime is upset. This is due to an incomplete compensation of the emf across the resistance of the barrier layer r_1. This limits the theoretical upper limit of the resistance. Furthermore, and this is important in practice, an increase in the value of R_{ob} above a certain value does not lead to further increase in the accuracy of the measurement for a given sensitivity of the measuring circuit.

Greater accuracy can be achieved if simultaneously with the increase in R_{ob} one increases by the same factor the sensitivity of the circuit. However, a decrease in R_{ob} is also inconvenient since the relative accuracy of measuring small luminous fluxes will decrease as a result, which is undesirable. For all these reasons the value of R_{ob} which was used in practice was chosen to be 100,000 Ω and was in the form of a special resistance coil.

In order to improve the conditions of work with the apparatus and to exclude possible sources of error a number of precautions were taken and are listed below.

1) Stabilization of the working current of potentiometer. A working current of 0.01 amp was usually employed in the low resistance potentiometer. This current is greater by a factor of 100 compared with the working current of the high resistance potentiometer. An increase in the current leads to a loss in the stability of the voltage of the supply battery. For this reason it is necessary to check the working current in the low resistance potentiometers more often than in the high resistance ones. This slows down the work somewhat. In order to improve the stabilization of the working current of the potentiometer we used alkaline accumulators having a capacity of 90 amp-hours and a voltage of 3.75 v. During experiments this battery was continuously charged by another 5v battery through a MKMS resistance box. This resistance box was used in a continuous adjustment of the charging current. The charging current was chosen so that the potentiometer current changed by as little as possible.

This type of supply ensured a sufficiently stable current so that it was enough to check it only once in 20-30 minutes. This made the work with the measuring setup much easier.

2) Stability of the zero position of the cross-wire of the galvanometer.

The true zero position of the cross-wire, or, in short, the zero of the galvanometer, is usually taken as that position of the cross-wire which the latter takes up when the galvanometer is on open circuit. The closure of the galvanometer circuit with the FESS-U covered up (see Fig.) leads to some deviation of the cross-wire from the true zero. This new position of the cross-wire, with the photocell covered up, may conditionally be called the "false" zero. The true zero, as well as the "false" zero, usually drifts in an irregular manner to a lesser or greater extent. It is not always possible to exclude such a galvanometer drift. It is, however, possible to reduce this drift and thus increase the accuracy of measurement. Such a reduction of the zero drift is particularly necessary in connection with the compensation of small photoelectric emf's. Potentiometric measurements of large potential differences across the resistance R_{ob} are relatively less effected by the zero drift. In order to increase the accuracy steps were taken to exclude from the measuring circuit any stray currents, particularly noticeable in wet weather. A number of methods of doing this and usually recommended for this purpose [16, 19, 20] were tried. The enclosure of the whole measuring system by a closed metallic screen, which acted as a low resistance shunt for stray currents, did not always improve experimental conditions. In our experiments the most convenient method was to place all the instruments (potentiometer, galvanometer, resistance box, keys and switches), as well as the supply sources (standard cell and accumulators), on ebonite stands. Such an arrangement worked without fail under different external conditions. In order to prevent the transmission of vibrations to the galvanometer suspension the galvanometer was supported on pieces of microporous rubber. A screened lead was attached to the galvanometer while the body of the galvanometer was grounded. The use of unscreened leads led to an increase in the vibration of the galvanometer suspension. Steps were taken to prevent possible nonuniform heating of the various parts of the apparatus and the connecting leads. In the absence of air convection in the laboratory, the vibration of the cross-wire did not exceed ± 0.3 mm on the scale. The zero drift due to external causes was also not observed. Some drift was observed after an accidental application of a large current to galvanometer, but this was due to a residual angular momentum of the suspension wire.

It is well known that such precautions, although they lower the zero drift, do not exclude it altogether. This is particularly noticeable during wet weather. Various methods have been used by different authors [20-24] to exclude automatically residual emf's associated with stray currents. As a result of preliminary experiments with the circuit shown in the above figure we adopted the method suggested by Wenner [24, 25].

The following procedure was employed in applying Wenner's method to our apparatus: a) first, an approximate compensation of the photocell was obtained by altering the resistance R_1; b) by means of a special reversing key (not shown in the figure) the current in the galvanometer circuit was reversed. Each change of the direction of the current was carried out quickly but the interval of time between opposite directions was roughly equal to the period of the galvanometer. At the same time the final compensation of the photoelectric emf was carried out and this was taken to be completed when no deflection was obtained on reversing the current through the galvanometer.

By this method it is possible to exclude residual emf's and to carry out measurement with an accuracy up to 1×10^{-7} to 1×10^{-8} v.

However, our circuit together with the FESS gives an accuracy of only 2×10^{-7} - 3×10^{-7} v. For this reason, and also because of a considerably decrease in the humidity in the laboratories of the Observatory, due to the introduction of central heating, it turned out (beginning with 1952) that this complex attachment did not have to be used and only those precautions listed above were taken.

We shall now consider the results of tests on the measuring circuit imcorporating a silver sulfide photocell. For constant work two photocells were used: FESS-2 No. 337 and FESS-U2No. 555 (Table 1). The FESS-2 No. 337 photocell was used up to April 1953 when it lost its sensitivity, apparently due to the penetration of a large external emf across the terminals of the photocell. Other FESS photocells, as well as selenium photocells, were tried in the above circuit. First of all, the direct proportionality between the photocurrent and the incident luminous flux was tested many times, over a number of years, for large changes in the luminous flux. The currents were so chosen that they corresponded to those usually met in spectrophotometric and pyrometric studies of black bodies and other sources of radiation. Measurements were carried out on a photometric bench with a precision three meter distance scale. An electric lamp with a conical spiral 4 mm in size, or a matt screen 2 mm in size and illuminated by this lamp from a small fixed distance were used as sources of light. The lamp was placed in a box with an aperture which transmitted light only in the direction of the photocell. In order to exclude stray scattered light a system of diaphragms covered with black velvet were used, and, in addition, directly in front of the photocell a blackened tube containing a number of diaphragms was placed. These steps ensured a complete exclusion of stray light and the photocell did not react when the directly transmitted beam was absent. The light source was controlled by means of a potentiometric scheme described in [18]. The source of light and the midpoint of the receiver of radiation were placed at the same height on the optical bench. All the above steps

TABLE 1

Characteristics of Silver Sulfide Photocells

Photocell designation	Photocell No.	Integral sensitivity	Photoelectric emf at 25 luxes	Date of preparation
FESS-2	337	6670	85	11/1948
FESS-U2	555	7700	98	10/1949

TABLE 2a

Results of an Investigation of the Linearity of the Scale of the FESS-U2 No. 555 Working on Short Circuit (see figure)

True distance of source from photocell (mm)	Reduction by Talbot's disc (+ with disc, –without)	Potentiometer readings	Observed relative reduction in incident flux	Mean reduction
500	–	57420	1	—
1000	» »	14236	4.03	—
500	» »	60560	1	—
1000	» »	15160	3.99	—
500	» »	60560	1	—
1000	» »	14973	4.04	—
500	» »	60560	1	—
1000	» »	15065	4.02	—
500	» »	—	—	1
1000	» »	—	—	4.02
1000	» »	14954	1	—
2000	» »	3755	3.98	—
1000	» »	14973	1	—
2000	» »	3715	4.03	—
1000	» »	14952	1	—
2000	» »	3736	4.00	—
1000	» »	—	—	1
2000	» »	—	—	4.00
500	» »	60560	1	—
2000	» »	3754	16.13	—
500	» »	60560	1	—
2000	» »	3756	16.12	—
500	» »	60560	1	—
2000	» »	3736	16.21	—
500	» »	60560	1	—
2000	» »	3741	16.19	—
500	» »	—	—	1
2000	» »	—	—	16.16
1000	» »	14083	1	—
1000	+	620	22.71	—
1000	–	14027	1	—
1000	+	619	22.66	—
1000	–	14038	1	—
1000	+	619	22.69	—
1000	–	—	—	+
1000	+	—	—	22.69
2000	–	3755	1	—
2000	+	165	22.76	—
2000	–	3736	1	—
2000	+	165	22.64	—
2000	–	3715	1	—
2000	+	164	22.65	—
2000	–	—	—	—
2000	+	—	—	22.68
1000	–	14953	1	—
2000	+	165	90.62	—
1000	–	14953	1	—
2000	+	164	91.18	—
1000	–	14973	1	—
2000	+	165	90.75	—
1000	–	14973	1	—
2000	+	164.5	91.02	—
1000	–	—	—	1
2000	+	—	—	90.89
500	–	60560	1	—
2000	+	165	367.0	—
500	–	60560	1	—
2000	+	164.5	368.1	—
500	–	60560	1	—
2000	+	165.5	365.9	—
500	–	—	—	1
2000	+	—	—	367.0

ensured a successful use of the apparatus in reducing the flux incident on the photoelement by a factor of 16-50 by varying the distance of the source from the receiver. Further reduction in the incident flux was obtained by means of a Talbot disc with a constant aperture which was prepared in the workshops of the Kuchin Observatory. The Talbot disc was placed near the light source. As a preliminary measure it was carefully tested in the Laboratory. The disc cut down the luminous flux by a factor of 22.77. In order to exclude possible changes in the sensitivity of the scheme during measurements all the measurements were repeated.

As an example, we shall give results of tests on the photocell FESS-U2 No. 555 using the circuit shown in the figure (Table 2a and 2b).

TABLE 2b

Results of an Investigation of the Linearity of the Scale of the FESS-U2 No. 555 Working on Short Circuit (see figure)

Theoretical reduction in the flux	Observed reduction in flux	Deviation (theoretical–observed)
1.00	1.00	–
4.00	4.01	– 0.01 (0.25 %)
16.00	16.16	– 0.16 (1.00 %)
22.77	22.68	+ 0.09 (0.4 %)
91.08	90.89	+ 0.19 (0.2 %)
364.30	367.00	– 2.7 (0.74 %)

The first column of Table 2a gives the distance between the source of the luminous flux and the photocell. The second column indicates whether Talbot's disc was used to reduce the incident flux. The third column shows the change of potential difference across the resistance R_{ob} measured by means of the potentiometer with photocell illuminated. This change in the potential difference, or the readings of the potentiometer, are proportional to the luminous flux incident on the photocell. The fourth column gives the relative reduction in the incident luminous flux obtained from the data given in the third column. The mean values for the reduction obtained from observations are given in the fifth column.

Table 2b gives a comparison of the theoretical and observed reduction in the luminous flux. The theoretical reduction was determined from the data given in the first and second columns of Table 2a.

It follows from Table 2a that the observed reduction in the luminous flux proceed by Talbot's disc is in agreement with the theoretical reduction within the limits of experimental error. This indicates that the reduction in the luminous flux by means of Talbot's disc can be used in the case of a photocell in combination with an inertial measuring instrument (galvanometer). Tables 2a and 2b show further that, in the case of a short circuited FESS in combination with a potentiometer (see figure), the deviation from direct proportionality between the photoelectric current and the magnitude of the incident flux does not on the average reach 1% and is equal to 0.5-0.7%. The apparatus considered here may be used to measure directly, and with the same sensitivity, luminous fluxes which differ in magnitude by a large factor. This important property of the apparatus is particularly necessary in absolute measurements.

In addition to better linearity of the photocell scale the apparatus considered here (see figure) lowers the temperature dependence of the sensitivity of the photocell.

Shternberg State Astronomical Institute

Received November 12, 1956

LITERATURE CITED

[1] G. F. Sitnik, Absolute Photoelectric Photometry of the Continuous Solar Spectrum [in Russian] (MGU, Moscow, 1955).

[2] G. F. Sitnik, Vestnik MGU, Ser. Fiz.-Mat. Nauk 1956, No. 1, 63.

[3] G. F. Sitnik, Doklady Akad. Nauk SSSR 110, 193, 780 (1956).

[4] V. E. Lashkarev, J. Tech. Phys. (USSR) 18, 1347 (1948).

[5] I. F. Potapenko, J. Tech. Phys. (USSR) 18, 1356 (1948).

[6] V. E. Kosenko and E. G. Miseliuk, J. Tech. Phys. (USSR) 18, 1369 (1948).

[7] E. G. Miseliuk, Silver Sulfide Photoelectric Cells FESS-U [in Russian] (Izd-vo In-ta Fiziki AN SSSR, 1949).

[8] A. A. Il'ina, Izv. AN SSSR, Ser. Fiz. 11, 436 (1947).

[9] V. E. Lashkarev, and K. M. Kosonogova, Doklady Akad. Nauk SSSR 54, 125 (1946).

[10] A. V. Markov, J. Tech. Phys. (USSR) 16, 1195 (1946).

[11] G. F. Sitnik, Astr. Tsirkuliar 1956, No. 166, 16.

[12] S. Iu. Luk'ianov, Photocells [in Russian] (Izd-vo. AN SSSR, 1948).

[13] L. A. Wood, Rev. Sci. Instr. 6, 196 (1935); 7, 157 (1936).

[14] Compball and Freeth, Journ. Sci. Instr. 11, 125 (1934).

[15] E. S. Rittner, Rev. Sci. Instr. 18, 36 (1947).

[16] V. I. Krotova, Potentiometers [in Russian] (Izd-vo. VNIIM,* Leningrad, 1940).

[17] Description and Directions. A Low Resistance DC Potentiometer of Type PN-2 [in Russian] ("Etalon" factory, 1947).

[18] G. F. Sitnik, Trudy GAISh** 22, 3 (1953).

[19] Krukowski, Arch. Tech. Messen 1935, No. 52, 117.

[20] Temperature, its Measurement and Control in Science and Industry (N. Y., 1941).

[21] H. Diesselhorst, Z. Instrumentenk. 28, 7 (1908).

[22] W. P. White, Phys. Rev. 25, 344 (1908).

[23] H. B. Brooks and A. W. Spinks, Bur. Standards J. Research 9, 781 (1932).

[24] F. Wenner, Bur. Standards J. Research 22, 425 (1939).

[25] R. P. Teele, Bur. Standards J. Research 22, 431 (1939).

---

*All Union Scientific Research Institute of Meteorology.

**Shternberg State Astronomical Institute.

CHANGES IN THE CONSTRUCTION OF THE MF-2 MICROPHOTOMETER ADAPTING IT TO MEASUREMENTS OF HIGH DENSITY

D. Ia. Martynov and V. Ia. Alduseva

As a result of the replacement of the valve photocell in the MF-2 microphotometer by a photomultiplier, and the use of a limiting slit which eliminates scattered light, it is possible to measure densities of the order of 3-4 or more with satisfactory accuracy, which makes it possible to study highly overexposed spectral lines along with normally exposed parts of the spectrum. The necessary condition for this is that the characteristic curve of the photoemulsion should have a sufficiently long straight part and that the range of the emulsion be large. Suggestions are made about the best arrangement of a step-scale.

We have removed from the MF-2 microphotometer the valve photocell having a sensitivity of 410 μa/lu and have replaced it by the FEU-10 photomultiplier which has a sensitivity of 22.8 amp/lu for an overall potential difference of 1250 volts. The linearity of the photomultiplier was checked on an optical bench. The optical bench was 13 m long and this ensured a verification of linearity over the required range of illumination. Four such photomultipliers were tested on the optical bench and the most satisfactory one was chosen.

After having been set up on the microphotometer the photomultiplier was used in conjunction with the PS-26 galvanometer having a sensitivity of $2.6 \cdot 10^{-9}$ amp/mm/m.

The galvanometer was shunted by a variable resistance in the form of a plug-in resistance box P-14. In addition, in order to increase the density range, neutral light filters whose absorption was accurately determined were used.

The exclusion of scattered light turned out to be quite a difficult problem. The illuminated slit of the MF-2 microphotometer which is made of two moveable green glass plates was unsuitable for our purposes. It turned out that the use of this slit introduced an error of the order of 10% already for densities of the order of 2. The slit was therefore replaced by a modified Hartmann diaphragm, so that the part of the plate illuminated by it was in the form of a narrow rectangle whose height and width could be altered. The resulting inconvenience of seeing only a small part of the negative at a time on the screen is inevitable but can be alleviated to some extent by fast removal of the diaphragm beyond the field of view if the negative as a whole is to be inspected.

After the above modifications the microphotometer must be used in a dark room.

The following procedure was adopted.

The HgI triplet ($\lambda\lambda$ 4339.24, 4347.50 and 4358.35) was photographed on an antihalation plate placed at the exit slit of the universal monochromator UM-2, the slit being sufficiently wide to accommodate all the three lines. The exposure was chosen so that even the less intense components of the triplet were over-exposed. The exposure was then repeated with a step wedge placed in front of the entrance slit. Finally, the same three lines

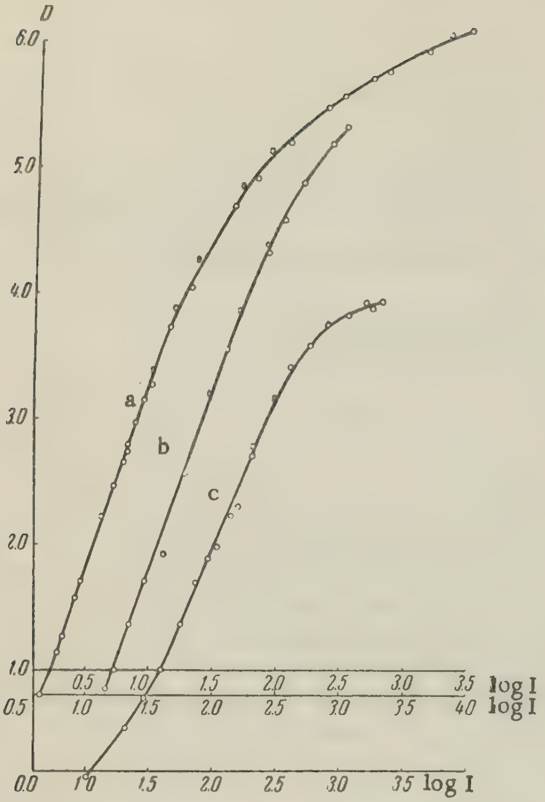


Fig. 1.

were measured on the monochromator directly, using the same photomultiplier which was used in the microphotometer (this is, of course, optional). During this measurement the very narrow exit slit of the monochromator was placed very accurately in the focal plane containing the image of the measured lines and their photometric profile was obtained by a micrometric displacement of a dispersing prism so that a given line could be moved slowly across the slit. The width of the entrance slit was kept constant throughout the above procedure. Normal slit width was used. The illumination conditions, on the other hand, could change since the mercury lamp used to illuminate the slit could be switched off.

The photometric profile of the lines obtained in this way was compared with the profile obtained from measurements on the photograph of the spectrum obtained in the first stage.

In Figure 1, Curve a shows the characteristic curve of the Agfa Ultra Finopan plate which was used to obtain the mercury lines. While the measurements which can be carried out on the modified microphotometer extend up to densities of the order of D = 6.0, the straight path of the characteristic curve which ensured a high and constant accuracy of photometering reaches up to D = 4.0.

The maximum densities obtained for the three mercury lines were the following:

λ	D
4358.35	5.8
4347.50	4.1
4339.24	3.1

The characteristic curve a was used to convert the measured densities into intensities after which the profiles of the three lines were plotted (Fig. 2) and compared with the profiles obtained by direct measurements on the monochromator (Fig. 3). A detailed examination of Figure 2 and Figure 3 shows that although they are generally similar, nevertheless there are quantitative differences between them, and, in particular, the differences between the logarithms of the maximum intensities of the lines are as follows:

	From the photograph	From the photoelectric measurements
log I 4358 − log I 4347	1.580	1.260
log I 4347 − log I 4339	0.400	0.420

It is interesting to note that the intensity of the brightest line λ 4358 relative to the minimum between the lines λ 4339 and λ 4347 was roughly the same in both cases, namely, 2.777 and 2.700, indicating that the error was not in the intensity of the brightest line but in that of the weaker line. This made the search for the error more difficult.

An analysis of the experimental conditions showed that the problem of constructing an accurate characteristic curve covering a large interval of densities, using the existing step wedge filters, can only be solved with difficulty. In fact the filters currently produced by our industry give intensity changes in the range 1 : 10 or 1 : 12 so that in our case the characteristic curve must be plotted in parts, using three or four sections and then making them fall on a single line by fixing the end points, an operation which is not free from subjective errors. Furthermore, even a small departure from parallelism in the slit may introduce an appreciable systematic error. In order to resolve these difficulties one would require a step wedge filter with an intensity range of 1:100 or even 1:1000.

A second source of error was the diaphragm which replaced the illuminating slit of the microphotometer, since it was not sufficiently accurately made and did not completely eliminate the scattered light, so that its width could not be kept constant during the measurements, which meant that the fraction of scattered light was variable. For this reason the diaphragm was replaced by a slit, specially prepared in the workshops of the GAISh, the width of which could always be reproduced.

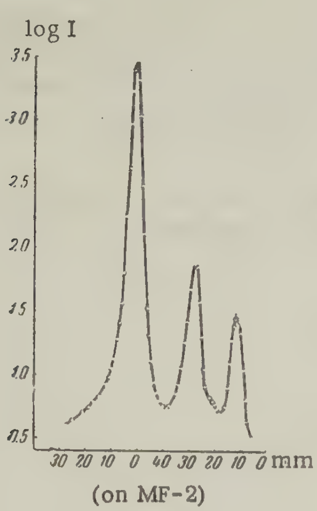


Fig. 2.

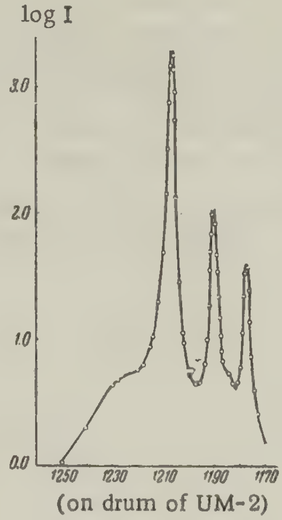


Fig. 3.

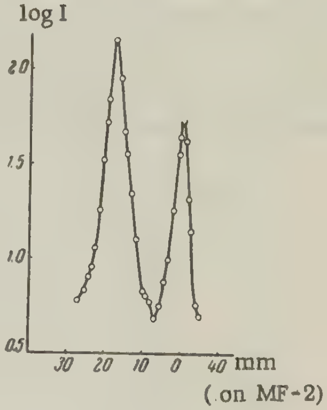


Fig. 4.

This modification soon produced results: when the measurements on the negative of the most intense line λ 4358 were repeated, the deflections of the galvanometer remained within the limits of the dark current fluctuations, while during the original measurements on the same negative the deflections were quite appreciable, i. e., they were due to scattered light.

As a result of taking into account the nonparallelism of the slit, and a new and very careful study of the step-wedge filter, a new characteristic curve was obtained for the same spectrogram of Agfa Ultra Finopan and is shown in Figure 1, b; this curve does not reach such high densities but its straight part is longer.

The photometric profiles of the lines λ 4347 and λ 4339 obtained from the new measurements are shown in Figure 4. As was stated above, the line λ 4358 turned out to be inadmissible. The difference $\log I_{4347} - \log I_{4339}$ turns out to be equal to 0.425 which is in excellent agreement with the result of photoelectric measurements, namely, 0.420. In actual fact, this result is no better than that obtained at first (0.400), but now we have a more correct value for the intensity of the two lines relative to the minimum between them i. e.,

λ	$I_{\text{photog.}}$	$I_{\text{photoel.}}$
4347	1.385	1.440
4339	1.060	1.025

The weakness of the above experiment lies in that the conditions of illumination of the slit in the case of the direct photoelectric measurement of the three lines of mercury and when they were photographed were not identical. As far as possible, all the experimental conditions were noted in the following series of experiments.

The lines were now photographed using Ilford Ordinary plates. Their characteristic curve is shown in Fig. 1c. Its linear part is shorter than that of Agfa Ultra Finopan (we did not have any more plates of the latter kind) but its latitude is smaller, as a result of which we could not obtain measureable densities in-between the mercury lines. Results of direct photoelectric measurements are shown graphically in Figure 5, while the measurements on the spectrogram are shown in Figure 6 and in the following table:

λ	D_{max}	$\log I_{max}$		$\Delta \log I_{max}$	
		photographic	photoelectric	photographic	photoelectric
4358	3.697	2.880	3.190	1.300	1.280
4347	0.996	1.580	1.910	0.475	0.440
4339	0.263	1.105	1.470	1.775	1.720

In a large number of cases, for example, in studying the spectra of explosions, the spectra of new stars, and other objects having bright lines, the observer meets with the difficulty of photometering all the details on a single plate, which is practically impossible since with normal exposure some of the details are hopelessly over-exposed and others are under-exposed. Furthermore, many photoemulsions have a sufficiently long linear part of their characteristic ranging up to densities of the order of 4-5, where the accuracy of conversion into intensities could be quite sufficient if only the microphotometer could give a sufficiently accurate measure of such densities. The latter is not, however, the case in the existing microphotometers.

The authors have come across this difficulty in connection with some measurements on strongly over-exposed lines of hydrogen and Ca II on the spectrograms of the solar eclipse on June 30th, 1954.*

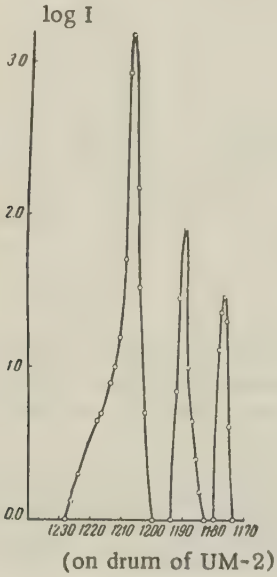


Fig. 5.

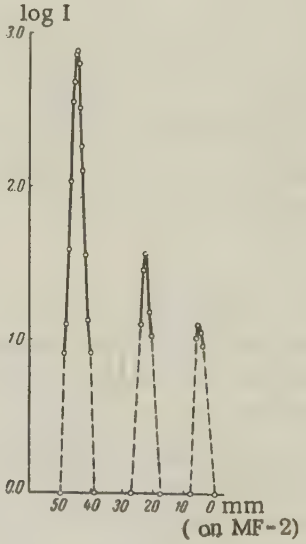


Fig. 6.

After a preliminary series of experiments it became clear that the problem may be solved only along the following two lines:

1. An increase in the sensitivity of the light receiver in the microphotometer;

2. Elimination of all traces of parasitic light scattered by parts of the negative neighboring the region under investigation.

The most convenient light receiver which will record light fluxes which change by a factor of 10^6 is the photomultiplier carefully checked for linearity. Of course, one cannot use a single indicating instrument under such conditions.

---

* The dilution of the negative, as suggested by T. V. Krat (Izvestiia GAO, 19, No. 151, 1953), which is suitable for our purposes will not be used for unique photographs since it leads to the destruction of part of the image.

These results as well as those obtained earlier show that the method described above can be used to measure densities up to 4-5, and to obtain intensities within accuracy of up to 10-12%. This accuracy is not very inferior to the accuracy of photographic photometry in the case of the more usual densities, and can probably be improved by using step-wedge filters, or other means of standardization of photographic plates, giving a wide calibration of illumination (up to 1:100 and 1:1000). Step-wedge filters of wide range should be made without opaque bands between the steps, while the steps themselves should be adjacent to each other, or possibly separated by semi-transparent bands of constant density if the steps can be made sufficiently wide (along the slit height).

The determination of photometric profiles of spectral lines is of course a long procedure if one uses the MF-2 microphotometer. However, the modifications made by us in the MF-2 microphotometer can also be made in the MF-4 microphotometer.

In addition, one needs an automatic switch-over of the galvanometer shunts so that its range is automatically adjusted. Our results for the MF-4 microphotometer will be reported later.

Shternberg State Astronomical Institute

Received November 12, 1956

ON THE IDENTIFICATION OF THE EMISSION OF THE NIGHT SKY AT ABOUT 5300 A

A. V. Mironov and V. S. Prokudina

The emission at about 5300 A evidently belongs to the vibrational spectrum of hydroxyl in the ground state. As this emission originates at a height of about 100 km it contradicts the ozone-hydrogen hypothesis and favors the oxygen-hydrogen mechanism of production of hydroxyl emission.

Koomen, Scolnik and Tousey [1] have shown, from rocket photometric measurements, that the diffuse emission at about 5300 A originates at heights between 80 and 115 km and has a maximum intensity at about 100 km. Earlier, Hunaerts and Nicolet [2] suggested that near 5300 A there should be the bands 6-0 and 9-2 of hydroxyl in the ground state. If this is the case, then the experiment of Koomen, Scolnik and Tousey determines the height at which the hydroxyl emission of the night sky originates and thus solves the problem of the origin of this emission. At the present time, two mechanisms of excitation of hydroxyl bands are being suggested. According to Bates and Nicolet [3, 4] excited hydroxyl OH^* appears as a result of the reaction

$$O_3 + H \rightarrow OH^* + O_2,$$

and this reaction cannot be effective above 80 km. According to V. I. Krasovskii [5] excited hydroxyl should appear as a result of a reaction between vibrationally excited re-formed molecules of oxygen in the zone of their effective recombination with hydrogen atoms according to

$$O_2^* + H \rightarrow OH^* + O,$$

and this reaction cannot be significant below 80 km while its maximum effect should be at a height of 100 km. Thus the experiment of Koomen, Scolnik and Tousey tend to refute the hypothesis of Bates and Nicolet and support the hypothesis of V. I. Krasovskii if the emission near 5300 A is in fact due to hydroxyl.

It is therefore important to identify the emission near 5300 A. Emission near 5300 A was found in one of the photographs in the archives of the Zvenigorod scientific station of the Institute of the Physics of the Atmosphere of the Academy of Sciences of the USSR. This photograph was obtained at the station on the 19th and 20th of March of this year using the SP-48[1] spectrograph with a slightly pre-exposed film RF-3 and an exposure of 14 hours[2]. The wavelengths of the emission found on the photographs are given in the table. The table also gives calculated wavelengths of the bands 6-0 and 9-2 of hydroxyl according to the data of Hunaerts and Nicolet and the suggested identification.

---

[1] Diffraction spectrograph SP-48 (GOI) aperture of the camera F: 0.8, dispersion 84 A/mm (λ 5500 A). The slit width was 0.2 mm. The spectrograph was directed northward at an angle of 25° to the horizon.

[2] It should be noted that the transparency of the atmosphere during these nights was rather good.

The band 6-0 should evidently be more intense than the band 9-2 since the former corresponds to the sequence $\Delta v = 6$ and the latter to the sequence $\Delta v = 7$. In reality the reverse is observed. It is possible that this result is due to some defects of the photograph.

λ Observed	λ Calculated	Band	Branch	Value
5181	5187.4	9—2	R_1	5/2
5192	5189.3	9—2	R_1	7/2
5202	5203.9	9—2	Q_1	3/2
	5209.1	9—2	Q_1	5/2
5214	5216.9	9—2	Q_1	7/2
5241	5239.3	9—2	P_1	7/2
5258	5254.5	6—0	R_1	7/2
5289	5287.2	6—0	Q_1	7/2
5305	5299.5	6—0	P_1	5/2
	5313·8	6—0	P_1	7/2

In general the above photograph supports the suggestion that the emission near 5300 A is due to hydroxyl. However, in order to verify this completely, it is necessary to repeat the photograph of the spectrum of the night sky near 5300 A, which we intend to do as soon as suitable meteorological conditions occur.

We take this opportunity to thank V. I. Krasovskii for valuable remarks during a discussion of the above results.

Zvenigorod Scientific Station of the
Institute of Physics of the Atmosphere
of the Academy of Sciences of the USSR

Received September 8, 1956

LITERATURE CITED

[1] M. J. Koomen, R. Scolnik and R. Tousey, Astron. J. 61, 182 (1955); O. E. Berg, M. Koomen, L. Meredith and R. Scolnik, J. Geophys. Research 61, 2 (1956); M. Koomen, R. Scolnik and R. Tousey, J. Geophys. Research 61, 2 (1956).

[2] J. Hunaerts and M. Nicolet, Ciel et Terre 66, 213 (1950).

[3] D. R. Bates and M. Nicolet, J. Geophys.Research 55, 301 (1950).

[4] D. R. Bates and B. L. Moisewitsch, J. Atmospherical and Terrest. Phys. 8, 305 (1956).

[5] V. I. Krasovskii, Uspekhi Fiz. Nauk 47, 493 (1952); 54, 469 (1954).

TABLES OF THE PRINCIPAL PLANETS

G. M. Clemence

In the present paper, which was written at the request of the editors of this journal, the fundamental astronomical tables, giving coordinates of the major planets, which are being issued at the present time as a result of the progress made in calculating techniques, are described, as well as the revision of the theory of motion of the major planets which is being carried out under the author's direction.

At the beginning of the present century, the tables of Newcomb and Hill (with the corrections for Mars given by Ross) became the main source for the ephemerides of the sun and planets for all the national astronomical year books except for Connaissance des Temps which continues to make use of the tables of Leverrier and Gaio. In order to establish the nature of the existing systematic differences and to find their source it would be of interest to have a detailed comparison between the ephemerides in the Connaissance des Temps and other national ephemerides. There is doubt as to whether such a comparison will ever be made. Astronomers who are able to do this prefer as a rule to improve the work of their predecessors by producing their own planetary theories instead of improving in detail the theories of others.

As from 1960 the tables of Newcomb and Hill for Jupitor, Saturn, Uranus and Neptune will be replaced by coordinates given in Volume 12 of the Astronomical Papers of the American Ephemeris which includes also the coordinates of Pluto. As is shown in the introduction of Volume 12, a comparison with observations indicates that these coordinates, which were obtained by numerical integration of the equations of motion, are much more accurate than the tables used before. In fact this increased accuracy justifies the practical transition to the newly calculated coordinates. However, the increase in numerical accuracy was accompanied by real losses in other directions. In order to obtain detailed information about the motion of a planet one must know much more than just the numerical solution of the equations of motion in the interval 1653 to 2060. It would be desirable to be able to separate out from each other the various perturbing forces and to establish the effect of each of them. For this purpose it is further desirable to have expressions for the coordinates in a more general form, namely, in the form of functions of time so that the coordinates could be calculated for any instant of time during at least a few thousands of years. It would be better still to have solutions of the equations of motion which would hold over a million or more years. It is not known at present whether such solutions can be obtained, but the existence of fast calculating machines leads us to hope that this can be done.

Numerical integration of the equations of motion should therefore be looked upon as a temporary method to be used only until a more satisfactory form of solution is obtained.

In all probability Newcomb's tables for Mercury and Venus will be sufficient for all the needs of astronomy for many years yet. The accuracy requirements are least stringent for these two planets, because due to the effect of phase and the apparent closeness to the sun they cannot be observed with the same accuracy as the other planets.

From the point of view of contemporary requirements the tables of Newcomb and Leverrier are insufficiently accurate for the sun. During a close approach of a small planet to the Earth, these tables cannot be used to calculate its position with the same accuracy with which it is observed. The eclipse of the sun cannot be predicted using these tables with the same accuracy with which it can be observed. In order to obtain the longitude of the sun from Newcomb's tables it is necessary to combine sixteen quantities (in units of 0″.01) twelve of which are obtained by interpolation. The error in the sum in separate cases could reach a few units even if the tables themselves were very accurate. It is necessary to have a theory of motion of the earth which would give the longitude of the sun with an accuracy of at least 0″.01.

By taking Ross's corrections into account Newcomb's tables are deprived of the least agreement with equations of motion. Ross determined the change in the eccentricity, inclinations, longitudes of the perihelion and node over a hundred years using observations and thus introduced an unnecessary empiricity into the ephemerides of Mars. Furthermore, as Newcomb points out in the introduction, a large number of second order perturbations is absent in the tables of Mars.

Since 1941, I have been concerned, with some interruptions, with the setting up of a new theory of the motion of Mars. The first order perturbations have already been published and the calculation of the second order perturbations was finished in 1954. This theory in its present state agrees much better with observations of Mars than any other theory. I have, however, not published the second order perturbations since a comparison with numerical integrations over a period of 30 years shows differences in the longitude in the orbit of up to 0″.05,i. e., considerably more than the accuracy which one tries to achieve in calculations using either of the two methods. Until the reason for these divergencies is found the new general theory cannot be looked upon as fully satisfactory.

Although new general theories are necessary for all the major planets except for Mercury and Venus it does not necessarily follow from this that new tables are also necessary. Tables of the motion of planets are only an auxiliary tool which makes easier the calculation of trigonometric series which express the departure from elliptical motion which is due to the gravitational effect of perturbing planets. When the existing tables were produced they were very suitable for this purpose. For example Newcomb's expression for the longitude of the sun contains about 116 trigonometric terms with arguments depending on the planets; they are given in seven tables, five of which have double entries.

The use of these tables involves a calculation of eight arguments, interpolation from tables, and the summation of seven obtained numbers. Without the tables it would be necessary to calculate 116 arguments as well as the sine and the cosine of each of them and then to multiply them by the corresponding coefficients and obtain 116 products. Clearly, in the absence of calculating machines tables are an important help to calculators. Even in the presence of slow machines tables are still useful. A few years ago, when it was decided to calculate the solar ephemeride for each fourth day from 1800 to 2000 it was found that it is about ten times as productive to record Newcomb's tables on punched cards, choosing the interval so that one could interpolate with second differences, than to calculate each of the 116 terms separately for the required date. However, the advantages of tables are not so great when fast electronic machines are available. However, such machines are not well adapted to memorizing very voluminous tables on which various operations have to be carried out and therefore in order to be able to use their high speed it is necessary to limit the initial data to a few thousand. Using an electronic calculating machine with instructions and numbers recorded on a magnetic drum, a series of 116 terms may be separately calculated and summed in about 20 seconds. This time may be considerably shortened if the calculation is well planned as is indeed done in preparing extensive calculations of ephemerides. In the case of the fastest machines available at present, a series of 116 terms may be calculated in less than a second. It is to be expected that even faster machines will be come available in the future.

The problem therefore arises whether one should spend time on preparing and printing new tables of planets as soon as new general theories appear. Would it not be sufficient to publish the trigonometrical expressions for the coordinates, with an ephemeride for one hundred or a few hundred years? Undoubtedly the difficulty involved in obtaining such an ephemeride is comparable with the difficulty of setting up a table. If that is the case, what should be provided for the calculation of the position of the planet in very distant epochs beyond the limits of the ephemerides? No definite answer can be given to these questions at present.

In conclusion of a discussion of planetary tables something should be said about solar tables which have a special position since they are used in the calculation of the geocentric position of each planet. It is well known that if the ecliptic is chosen as the plane to which the planes of orbits of the planets are referred, the number of elements of the earth's orbit decreases from six to four. The inclination and longitude of the node is absent in solar tables; these tables give the longitude and latitude of the sun relative to a moving ecliptic, while the tables of the remaining planets take into account the effect of motion of the ecliptic.

However, it is not so well known that the new determination of the second as a unit of time has reduced the number of the elements of the earth's orbit from four to three. The second is now defined as a certain fraction of the tropical year 1900.0. The numerical value of this fraction was obtained from the coefficient of T in the expression for the mean longitude of the sun according to Newcomb's tables, by dividing it by 4,089,864,960,000,000 (product of the numbers: 1,296,000 – number of seconds of arc in a circle, 36,525 – number of days in a Julian year, and 86,400 – the number of seconds in a day); Newcomb thought that the coefficient of T was an element of the earth's orbit which must be determined from observation and this was the case as long as the unit of time was the mean solar day. However, by defining the second through the value of the coefficient of T which was given by Newcomb, we take this coefficient as an absolute constant, which determines the scale and the unit of time, and need not be considered any further.

Thus if we wish to avoid a discontinuity in the measurement of ephemeroidal time, the constant term in Newcomb's expression for the mean longitude of the sun should also be looked upon as an absolute constant which establishes a fundamental starting epoch for ephemeroidal time. Thus the number of elements of the earth's orbit decreases from three to two. It may turn out that at some time after a few centuries, when the difference between ephemeroidal and mean solar time will be much greater than now, it will be desirable to change this constant in order to reduce this difference. However, such a change will only be a matter of convenience and it will not be its purpose to bring the measurements of ephemeroidal time into agreement with observations. On the basis of the above we may consider that the earth's orbit has only two elements, namely eccentricity and longitude of the perihelion.

U. S. Naval Observatory, Washington

Received October 12, 1956

THE CHARACTERISTICS OF CHEMICAL COMPOSITION OF METERORITIC MATTER AND THE ORIGIN OF METEORITES

A. A. Iavnel'

Of special importance for solving the problem of the origin of meteorites are the characteristics of their chemical composition, which can throw light on the processes which took place in the early stages of development of meteorite matter. Prior, Chirvinskii, Brown and Patterson, Urey and Craig studied some of these characteristics, using incomplete data, and came to different conclusions about the origin of meteorites.

The author used data on meteorites of all three classes: stony, stone-iron and iron, and studied their phase composition — the metal and silicate separately. All the meteorites, depending on the relation between their phases, are divided into six subclasses: Ca-rich achondrites, Ca-poor achondrites, chondrites, mesosiderites, pallasites, siderites; according to their phase composition (the contents of Ni in the metal and that of FeO in the silicate) — into five groups.

If the different groups of meteorites of a given subclass are compared, a relation is found between the content and the quantitative phase relation, which is called Prior's "group" law. The author comes to the conclusion that every group of meteorites having a given phase composition was formed from a separate asteroid with layered structure. The relation between the content and quantity of each phase in different meteorites — Prior's "primary" law — is explained by the differentiation of matter (according to the melting points of the elements), which took place before the formation of separate asteroids. The oxidation of the elements which followed was the cause of the formation of silicate and metal phases. In each asteroid owing to phase differentiation under the influence of gravitational forces, separate layers, corresponding to the different subclasses of meteorites, were formed.

Among the various problems of cosmogony, the problem of the origin of meteorites occupies a definite place.

Meteorites are at present the only extraterrestrial bodies which can be actually handled by an investigator. Therefore the problem of their origin can be attacked not only by the methods of astronomy but also by those of geochemistry, mineralogy, petrography, metallography and other "earth" sciences.

As the result of such investigations, it is definitely known that meteorites belong to the solar system. According to the data of astronomy, the computed orbits of some meteorites and bolides are very similar to the orbits of asteroids. The resemblance between meteorites and terrestrial planets lies in the similarity of the isotopic composition of a number of elements found in meteorites and in terrestrial rocks. Moreover, determination of the "absolute age" of meteorites give 4.5 to 5 billion years, showing that meteoritic matter was formed contemporaneously with our planetary system, as is to be expected from astronomical evidence.

Thus, the problem of the origin of meteorites belongs to that part of the cosomgony of the solar system which deals with the region between the orbits of Mars and Jupiter.

At present there is a great deal of factual material on the composition of meteorites, which permits certain generalizations and provides an approach to the problem of their formation. Of especial importance is the study of variation in the chemical composition of meteorites, for it gives an insight into the processes which were active in the early stages of development of meteoritic matter and provides a key to its later history.

A brief review of the main results of the few investigations devoted to the study of the chemical characteristics of meteorites is given below.

Meteorites are mixtures of silicate (magnesium silicates) and metal (iron-nickel) phases. The relative proportions of these phases make it possible to group meteorites into three classes — stones, stony-irons and irons. It should be noted that studies were usually made of composition characteristics of each class separately and especially of the stony meteorites, which are subdivided into chondrites, consisting largely of minute globular structures, chondrules, and achondrites, which do not exhibit these structures.

One of these studies is that by Prior [1], who established the following rule on the basis of analyses of stony meteorites: "The smaller the concentration of the metal phase in a chondrite (italics mine, A. Ia.), the greater the amount of nickel in the metal and the greater the concentration of combined iron in the silicate phase." * Deviations from this rule Prior ascribed to faulty analyses. He believed that this variation in composition is continuous, but for the sake of classification divided chondrites into four groups according to composition, later reducing these to three [2].

The chondrites constitute nearly 9/10 of all stony meteorites, and Prior believed that his rule extended to other stony meteorites, achondrites, which contain less iron-nickel than chondrites, and also to stone-irons, which contain more. In that case, achondrites should be expected to contain iron-nickel and silicates in order to be correspondingly enriched in nickel and iron and the stony irons be less rich. Actually this is not true, but Prior dismissed the contradiction as an "exception which proves the rule."

On the basis of his generalization, Prior deduced that meteorites of different composition were formed from the same magma, as the result of progressive oxidation of iron-nickel. He assumed that the magma contained enough oxygen to form oxides of silicon, magnesium, aluminum and other elements exclusive of iron. With further oxidation, some of the iron was also oxidized; the nickel contained in it, however, was not oxidized, but became concentrated in the remaining iron. According to this, there must exist a continuous series of meteorites with gradually decreasing metal phase, from stone-iron meteorites (containing on the average 50% of metal) to achondrites (with less than 3% metal) and with corresponding gradual change in the composition of the phases.

Later, Prior [2], having studied the composition of the phases in stony meteorites, changed his earlier statement to read: "In any stony meteorite (italics mine, A. Ia.) the greater the content of nickel in iron-nickel, the richer are the magnesium silicates in ferrous oxide." * However, it is not difficult to see that although the new relationship is broader than the old (for all stony meteorites instead of for chondrites only) nothing is said here of the relation between composition and amount of each phase, the relation which Prior found in chondrites.

By this statement Prior admitted, in effect, the absence of this relation in stony meteorites (both chondrites and achondrites), i. e., of a continuous change in the composition of the phases with decrease in iron-nickel from chondrites to achondrites. Prior's proof of the origin of meteorites from a single magma thus lost an important argument, but was retained nevertheless.

Considerable significance was asigned by Spencer [3] to the rule found by Prior. The second rule (for stony meteorites) he named "Prior's law". From then on, this name has often been used in the literature, although most authors usually mean by it, the first rule (for chondrites), and sometimes cite it inaccurately.

---

* Translated from Russian; quotation not verified — Publisher's note.

P. N. Chirvinskii devoted much study to the clarification of the principles of chemical and mineralogical composition of meteorites and published his results in 1939 [4]. He also set a high value on "Prior's law" and gave slightly modified quantitative relations of components in different types of chondrites.

In supporting Prior's explanation of his law, Chirvinskii advanced a thermochemical basis for it and pointed out that chondrites rich in iron-nickel are heavier, more refractory, and probably of more deep-seated origin than chondrites which are poor in iron-nickel. Chirvinskii formulated a number of other "laws" for chondrites:

I. "The law of equal molecular content of the Mg-Fe orthosilicate (olivine) and the Mg-Fe metasilicate (bronzite, clinobronzite) in an average chondrite," means that an average chondrite contains 50 mole% M_2SiO_4 and 50 mole% $MSiO_3$.

II. "The law of identical distrubution of iron and magnesium in the orthosilicate and the metasilicate in an average chondrite," means that the following molecular proportions are observed in chondrites: for olivine — $5Mg_2SiO_4 : 2Fe_2SiO_4$, for pyroxene — $5\ MgSiO_3 : 2FeSiO_3$.

III. "Avogadro's law for the pseudoelement chondrium" states that the number of atoms per 1 cm^3 of chondrium is $2.014 \cdot 10^{23}$ and its atomic weight, 24.36.

It must be pointed out at once that Chirvinskii's concept of "pseudoelements," in the opinion of the present author, has no physical meaning, and therefore this "law" cannot be taken seriously.

IV. "The law of constant sulfur content (in troilite) with varying content of iron-nickel and oxygen" means that the content of sulfur in chondrites does not depend on the amount of iron-nickel, the content of nickel in the metal, or on the amount of oxygen.

In reality, it follows from Chirvinskii's data that the chondrites with the highest content of the metal component have a slightly higher content of sulfur than do those with lower content of metal.

Among other works, the paper by Henderson and Perry [5] should be mentioned, in which an attempt was made to compare the composition of iron meteorites and of the iron component of stony meteorites and also to present graphically Prior's rule for achondrites. However, because of insufficient data these authors could not arrive at definite conclusions.

These problems were investigated in greater detail by Brown and Patterson [6]. In their work the composition of both stones and irons was reviewed and hypotheses advanced concerning the structure of the planet which gave rise to meteorites. As the result of a statistical study of analyses, these authors determined the average content of the main components in the silicate and metal phases in stones and irons, and computed the spread of data. The average content of Ni (and Co) in the metal phase of the stones was found to be greater than in irons, due to enrichment of nickel in the metal phase and its decrease in amount in stony meteorites. It was noted that with a small content of metal the concentration of nickel in it is nearly double the average concentration of Ni in irons, while with metal content of 13% it is only slightly lower than this average.

Brown and Patterson believed that these results confirmed Prior's rule (for chondrites) which they cited inexactly by substituting in it the term "stony meteorite" for "chondrite," thus changing its sense. From their study of the gross composition of stony meteorites as a function of the amount of the metal phase present, the authors concluded that this relation is complex, showing inflections in its graph in the region corresponding to iron-nickel content of about 3%.

In reality these inflections are due to the mixing, by the authors, of all kinds of stony meteorites, chondrites and achondrites, which cannot be considered members of the same group with a gradual change in iron-nickel content. If only the chondrites are taken, the inflections disappear. Sulfur, in this case, shows the same relation as that given by Chirvinskii ("law IV").

Their data led Brown and Patterson to the conclusion that the state of oxidation of meteorites increases with decrease in metal phase and total iron content. Unlike Prior, who believed that meteorites separated from a single magma during its progressive oxidation, Brown and Patterson advanced the view that the differences in content of iron-nickel were due to gravity sorting within a large planet. It was difficult, however, to find reason for a rapid decrease in combined iron (FeO and FeS) with increase in the metal phase. The variation in the ratio Fe_{comb}/Fe_{met} was explained by the effect of pressure which shifted the equilibrium between these substances. The curve showing Fe_{comb}/Fe_{met} as a function of the iron-nickel content in stony meteorites has discontinuities at 4 and 14% of the metal phase. On the other hand, on the histogram of the metal phase content in stony meteorites, there are three maxima at 0-1%, 7-10% and 17-19%, and the troughs between these maxima coincide with the discontinuities mentioned above. The presence of these maxima were ascribed by the authors to seismic discontinuities of the second order such as are known on the earth.

Assuming an equilibrium between iron and nickel in the silicate and metal phases of meteorites, Brown and Patterson arrived at certain conclusions concerning temperature and pressure within the planet, which consisted, according to them, of a metal core and a silicate mantle.

The basic assumptions and results of this work were criticized by Klotz [7] and Levin [8], and are refuted by the results obtained by Urey and Craig [9], who used a different approach to explain variations in composition in meteorites.

Urey and Craig restricted themselves to stony meteorites. From the known analyses of these (about 350), they selected about 140 of the most reliable. After citing "Prior's law" (for chondrites), the authors pointed out that contrary to Prior's opinion the primitive meteoritic material was completely oxidized in the beginning, and was then subjected to reduction by hydrogen and oxidation by water, leading to variations in iron-nickel content.

Having reviewed the relation between Fe_{met} + Fe_{sulf} and Fe_{sil} in chondrites, they came to the conclusion that there are two groups of chondrites, which differ in total iron content. They showed, also, that one group cannot be derived from the other by simple oxidation and reduction or by mixing.

Concerning "Prior's law" (for chondrites), Urey and Craig expressed the supposition that the observed relation between the amount of iron in the metal and silicate phases of chondrites is due to faulty analyses. At the same time they affirmed the truth of this "law" in regard to the nickel content in the metal phase.

They pointed out a certain resemblance between their two groups of chondrites and Prior's groups, and set up special groups of enstatite, carbon and "amphoteric" chondrites, which differ from the latter in composition. Among the achondrites, Urey and Craig, like Prior, distinguished two groups which they called "chondritic" (Ca-poor according to Prior) and "basaltic" (Ca-rich) achondrites.

Considering the difference in composition of the two groups of chondrites, Urey and Craig conclude that meteorites were formed from two different asteroids. In their opinion, iron-nickel was melted out of the surficial layer of the asteroids and as the result beneath this layer, which gave rise to achondrites, a layer of FeS and a thin layer (a few meters in thickness) of metal were formed which eventually became the source of metal meteorites. Under the metal layer the main material of the asteroids remained unaltered — such, according to Urey and Craig, is the structure of asteroids.

These authors also expressed their views on formation of chondrites, which will be discussed later.

The results of Urey and Craig's work show the inadequacy of Prior's hypothesis of formation of meteorites from a single magma and point once more to certain regularities in variation of composition in chondrites. However in order to arrive at definite conclusions concerning the origin of meteorites, investigations cannot be confined to just one type of meteorite such as chondrites.

In his investigation of variations in composition of meteorites, the author of this paper used the most complete data on meteorites of all three classes, and studied, not their gross composition, but the composition of the metal and silicate phases separately without averaging these. This procedure made it possible to discover a number of new facts and to arrive at more definite conclusions concerning the early stage in the formation of meteoritic material as compared with previous investigators.

Many of these results were published by the author in earlier papers [10, 11] and they will be reviewed here briefly.

As stated before meteorites can be divided into stones, stone-irons and irons on the basis of the relative proportions of the metal and silicate phases. Moreover, further subdivision on the same basis is possible in the first two groups. Stony meteorites may be subdivided according to increase in the metallic phase into calcium-rich achondrites, calcium-poor achondrites, and chondrites; and stone-iron meteorites, into mesosiderites and pallasites.

Each of these subdivisions, which the author calls subclasses, has a characteristic structure which will be considered later. At present, then, six (together with iron meteorites or siderites) subclasses of meteorites may be distinguished.

Study of the composition of the main phases of each subclass shows that on this basis all meteorites can be placed in five groups. The division of meteorites into groups and subclasses is the basis of classification of meteorites proposed by the author [11].

The existence of groups is made especially clear by examining nickel content in the metal phase and FeO content in the silicate phase. Histograms of the composition of phases in meteorites of different subclasses reveal maxima or sufficiently isolated regions corresponding to definite meteorite groups.

It was discovered that the metal phase of iron meteorites of group I contains 5.6% Ni; of group II, 8% Ni on average; of group III, about 14% Ni; of group IV, 20-40% Ni; and of group V, over 40% Ni.

Examination of meteorites of the remaining subclasses showed that according to the composition of iron-nickel present they also can be referred to appropriate groups.

Thus among the pallasites were found representatives of groups I and III. The average composition of pallasites in each of these groups differs also in concentration of FeO in the silicate phase and in relative proportion of the phases, namely: pallasites of group III, which contain more Ni in the metal phase than pallasites of group II, contain more FeO in the silicate phase and have a smaller amount of the metal phase. Comparison of pallasites of different groups reveals a regularity analogous to that discovered by Prior in chondrites.

A similar regularity was found as the result of examination of phase composition in mesosiderites, which occur in group I but mainly in group II.

The most revealing results were obtained from the study of chondrites. In the author's works cited above, chondrites were referred to the first four groups. Chondrites of the Soko-Banja type and also carbon chondrites were assigned to group IV. Although according to the composition of the silicate phase the latter should probably be placed in group V, the composition of their metal phase calculated from the analyses then available placed them in group IV. However, according to the new analyses by Wiik [12], carbon chondrites contain considerably less than 4% iron-nickel (the amount given in the old analyses), so that Wiik could not even separate it chemically from his samples. Petrographic studies show that iron-nickel is present, rarely, in very minute grains and its content evidently does not exceed 2%. Considering that nickel in meteorites is concentrated in the metal phase and that its content in carbon chondrites is 1.4% on average [12], it may be assumed with some certainty that the concentration of Ni in iron-nickel of carbon chondrites amounts to no less than 50%.

It follows that according to the composition of both the metal and the silicate phases, carbon chondrites should be placed in group V.

It appears, then, that chondrites occur in all five groups. Those of the first four groups are similar to the chondrites of Prior's groups [1]. The two groups of chondrites separated by Urey and Craig correspond essentially to the most densely populated groups I and II, and the special groups of these authors (with the exception of "amphoteric" chondrites, which are really achondrites), also fit into this scheme of classification. This shows that Urey and Craig's groups of chondrites are only special cases in the subdivision of meteorites into groups and subclasses.

A study of the relationship between the FeO content in the silicate phase of chondrites and the amount and composition of iron-nickel revealed new characteristics.

First of all, it was confirmed that in passing from group I to group V the content of nickel in the metal phase of chondrites increases as does the content of FeO and of combined iron in general in the silicate phase. Very clear also is the connection between the amount of the metal phase in chondrites and the content of nickel in it, as shown in Fig. 1.

However, the graph shows that this relation is not of a continuous nature as supposed by Prior but discontinuous because of the presence of distinct groups of meteorites. This relationship is observed only in comparing chondrites of different groups and is absent within each group. For this reason this new relation found in chondrites, mesosiderites, and pallasites, the author calls "group" relation of Prior.

Examination of the composition of phases in calcium-poor and calcium-rich achondrites shows that they also can be placed in definite groups. When data were available on the composition of the metal phase of these meteorites they were placed in appropriate groups.

All this indicates a certain genetic relationship among meteorites of different subclasses which according to the composition of their phases belong to the same group. Because of the sharp distinction existing between individual groups it is natural to suppose that each group of meteorites formed from sufficiently individualized bodies of complex structure. Considering astronomical data it may be supposed that these bodies were separate asteroids.

The study of the variations in the composition of meteorites leads to the conclusion that the types of meteorites known at present were formed from five asteroids of slightly different composition and structure.

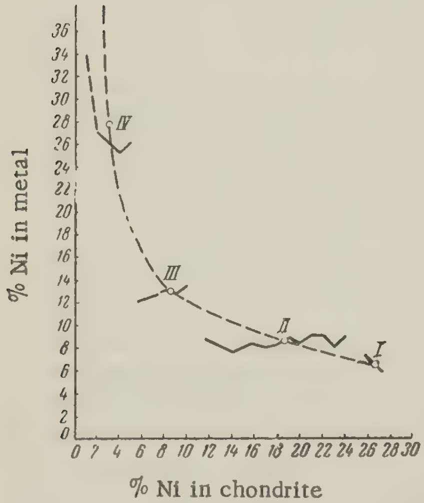


Fig. 1. Relation between the composition of iron-nickel and the amount of metal in chondrites.

This conclusion differs from that of Urey and Craig, who believed that meteorites were formed from two asteroids although they admitted the possibility that certain "special" types of chondrites resulted from a collision of these asteroids with other bodies.

The reasons for the difference between the present author's conclusion and Brown and Patterson's, who derived all meteorites from a single planet, will be considered in some detail. This difference is due mainly to the difference in interpretation of data. At the same time, these data find a simple explanation from our results.

For instance, the higher average content of nickel in the metal phase of stony meteorites as compared with iron meteorites is explained by different proportion of stony and iron meteorites in different groups. Stony meteorites are concentrated in groups II and III, while iron ones are concentrated in group II, which has a lower content of nickel in the metal phase, and this is reflected in the average composition of the metal phase.

Brown and Patterson's results show that stony meteorites with a low content of metal have nickel content twice as great as that in iron meteorites because they mixed chondrites and achondrites together. Had they treated chondrites separately, they would have obtained the result given here, i. e., a much higher nickel content in the metal of chondrites of groups IV and V as shown in Figure 1.

They concluded that in stony meteorites with 13% of metal and more the composition of iron-nickel is similar to that in iron meteorites, and this apparently suggested their common origin from one planet. This fact is also explained by Figure 1, which shows that stony meteorites containing 12 to 25% of iron-nickel are chondrites of group II which has 8% of nickel in the metal phase. This content is actually somewhat lower than the average concentration of nickel in iron meteorites, but as is easily seen, it bears no relation to it. In general, it must be said that the average composition of meteorites of any subclass — iron meteorites, pallasites, chondrites, etc., is not a genetic characteristic because it depends on distribution of meteorites among different groups as was shown above.

The maxima of iron-nickel in stony meteorites on Brown and Patterson's histogram correspond to our groups, while the relatively large number of meteorites which give a maximum at 0-1% of the metal phase is due to the fact that these authors combined chondrites and achondrites. Since the presence of natural groups among meteorites is explained by their formation from different bodies, it is not suprising that discontinuities occur in their graph "Fe_{comb}/Fe_{met} — amount of metallic phase."

In summary, there are no contradictions between our conclusions and the data on the composition of meteorites obtained by other authors.

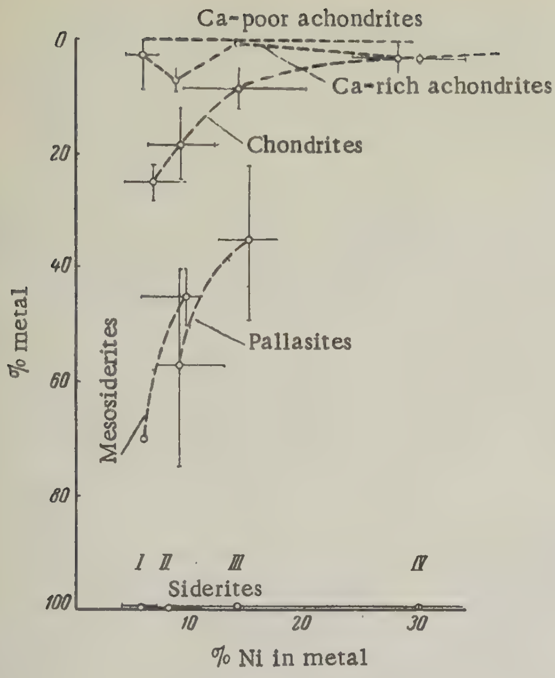


Fig. 2. Relation between nickel content in the metal phase and the amount of metal phase in meteorites of different subclasses.

The question of the structure of the asteroids which produced meteorites will be considered now. Definite clues are obtained from comparison of different subclasses of meteorites.

As has already been shown, according to the increase in the metal phase, meteorites are divisible into six subclasses, Ca-poor achondrites, Ca-rich achondrites, chondrites, mesosiderites, pallasites and siderites (iron meteorites). This sequence of subclasses persists within each group as can be seen from Figure 2.

The segments of straight lines on the diagram indicate maximum spread of values plotted on the coordinates; in reality this spread is considerably smaller. The absence of some subclasses in certain groups (for example, of pallasites in groups I, IV and V, mesosiderites in groups III – V, etc.) evidently indicates merely that meteorites of these types are rare and have not been found yet, not that their existence is impossible.

The continuity of the sequence of subclasses is confirmed by the presence of a small number of meteorites of types intermediate between adjacent subclasses, for example, between chondrites and Ca-poor achondrites, or pallasites and siderites, etc.

Thus in each group corresponding to an asteroid a regular variation with several maxima is observed in the relation between the metal and silicate phases. This can be explained by a layered structure of the asteroids resulting from differentiation. Most likely this differentiation was induced by gravity.

Considering differences in density of different layers, it may be assumed that the surficial parts of each asteroid produced Ca-rich achondrites containing minimum iron-nickel while the central parts (cores) produced siderites. All other layers must have been distributed between these two in the same sequence as the subclasses of meteorites.

It should be noted that the idea of layered structure of the parent body which produced meteorites is not new and that many authors believed that the body consisted of an iron-nickel core enveloped in a silicate mantle. Brown and Patterson even made an attempt to compute the size of such a body.

At the same time, as has already been pointed out, Urey and Craig arrived at a different idea of the structure of asteroids. In their opinion the silicate layer with a small content of iron-nickel was underlain by a thin layer of FeS and iron-nickel and then by a silicate core (corresponding to achondrites in composition?). The existence of a sulfide layer, which was probably introduced by these authors under the influence of Goldschmidt's idea of the internal structure of the earth, is doubtful. If such a layer really existed, meteorites composed wholly of iron sulfide should occur. However, FeS occurs in meteorites only as scattered small inclusions of the mineral troilite.

An especial peculiarity of Urey and Craig's conclusions is the place which they assign to iron-nickel in asteroids. It must be said that here their views contradict not only the established regularity in the subclasses of meteorites but also other facts.

If the iron-nickel layer, as these authors believe, had formed by accumulation of metal fused from the surficial layer of the asteroid, then achondrites should occur containing iron-nickel in a wide variety of concentrations. Achondrites with metal content over 10% are not known. Further, the thin layer of iron-nickel underlying the silicate layer only a few tens of meters thich (as supposed by Urey and Craig) could not, because of rapid cooling, develop coarsely crystalline texture found in iron meteorites. In relating the size of these meteorites to the thickness of the metal layer, the authors evidently believed that the texture of iron meteorites was formed under these conditions.

Other objections to this point could be found, but it is evident already that Urey and Craig's hypothesis of the structure of asteroids is subject to serious doubts since it contradicts a number of facts concerning the composition of meteorites.

On the basis of available data on the composition of meteoritic matter a few other questions connected with the origin of meteorites can be clarified. For this purpose the relation between amounts of the metal and silicate phases in different groups of meteorites must be examined.

Evidently this relation can be determined if for a given group the relation between different subclasses and their phase composition is known. So far as the latter is concerned, it is well illustrated by Figure 2, which also shows the regular variation in the amount of the metal phase in chondrites, mesosiderites, and pallasites in passing from one group to another (Prior's "group" variation).

Because of the small number of achondrites, mesosiderites and pallasites as compared to the number of chondrites and siderites, for an approximate evaluation of phase relation the last two subclasses will suffice. Indeed, the silicate phase of meteorites is concentrated essentially in chondrites, while the metal phase is found both in chondrites and iron meteorites.

In chondrites the determination of the ratio of the metal phase (M_{ch}) to the silicate phase (S_{ch}) is not difficult since the chemical composition of each group of chondrites is known. Its values are given in Figure 3, where the ratio M/S and the percent of Ni in the metal phase are plotted for each of the meteorite groups.

For computation of the ratio of metal phase, M_i, in iron meteorites to the silicate phase, S_{ch}, in chondrites the relative proportion of these subclasses must be known and this is determined from the number of falls and respective masses of these meteorites. Different investigators have estimated the ratio of iron meteorites to chondrites as being from 1:5 to 1:20. Taking the average ratio, 1:10, and knowing the number of falls and the phase composition of each type, the corresponding value of M_i/S_{ch} for different groups can be determined and plotted as in Figure 3.

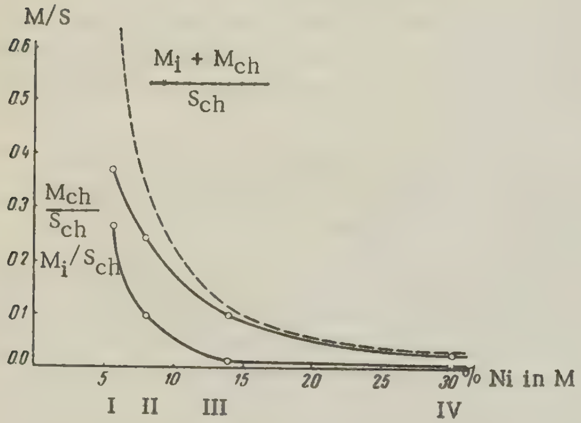


Fig. 3. Ratio of the amount of the metal phase (M) to the silicate phase (S) in different meteorite groups.

The value of M_i/S_{ch} decreases from group I to group V analagously with the ratio M_{ch}/S_{ch}. Evidently this relation is due basically to a certain ratio of the number of iron meteorites and chondrites in each group and will not be changed by using other average ratios than 1:10. The resultant of the two curves $\frac{M_i + M_{ch}}{S_{ch}}$, shown in Figure 3 by an interrupted line, represents the value of average proportion of the metal and silicate phases in different meteorite groups, i. e., in different asteroids. Assuming that before differentiation the material of each asteroid was relatively uniform, it must be concluded that this average proportion of the phases is the original one.

This throws light on the process of differentiation; namely the greater the initial relative amount of the metal phase the greater amount of this phase will pass into the iron-nickel core, as can be seen from Figure 3. At the same time, the relative content of the metal phase remaining in the different layers of the asteroid mixed with the silicate phase will also be the highest. This is the situation observed during fractioning of a heterogeneous mechanical mixture under the influence of gravity.

Thus, the so-called "group" variation of Prior, observed in different meteorite subclasses, is explained by phase differentiation in separate asteroids with different amounts of composition of the phases.

These initial values are connected with each other by a relation which by analogy may be named Prior's "primary" relation: the higher the content of metal in an asteroid, the poorer it is in nickel and the smaller is the amount of combined iron (FeO, FeS) in the silicate phase.

The question arises, how could this relationship have come about?

To answer this question it is necessary to examine the elemental composition of different asteroids, i. e., meteorite groups.

To this end, from the analyses of meteorites of each group the average content of the main elements, Fe, Ni, Si, Mg, O, S, C and H, totaling 98%, were calculated. The results of these calculations are shown in Figure 4, where, besides the total content of iron (Fe), its distrubution between iron-nickel (Fe_{Ni}), ferrous oxide (Fe_O) and iron sulfide (Fe_S), is also shown.

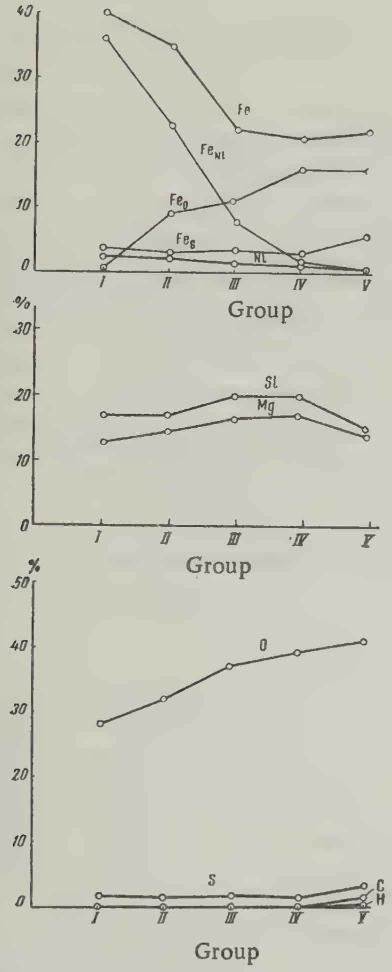


Fig. 4. Content of the main elements in different meteorite groups (asteroids).

It should be noted that the concentrations of some elements are approximate, especially in groups IV and V, because of inadequacy of data, but nevertheless their relative variation is quite definite.

The graphs show that the character of variation in content of the elements changes from group I to group V. The most refractory elements, Fe and Ni, have highest concentrations in the first group, and then it diminishes. The opposite is true of oxygen, whose content increases continuously, reaching maximum in group V. In this group also the content of sulfur, carbon and hydrogen increases sharply. The content of silicon and magnesium, elements with melting and boiling points intermediate between those of iron and oxygen, has a small maximum in groups II and IV.

This regular variation in concentration of different elements points definitely to an initial differentiation of matter while still in uncombined atomic state (possibly with the exception of carbon combined, most likely, with hydrogen), and before its separation into groups, i. e., before the formation of asteroids.

As the result of this differentiation of matter, the asteroids formed in different zones of condensation of proto-asteroid material differed in composition. As the material became denser (either just before separating into asteroids or after), molecules of different compounds were formed in it and first of all, oxides and sulfides. Later, these compounds combined to form minerals of the silicate phase of the asteroids.

Iron, nickel, copper and other metals with less affinity for oxygen than Mg and Si, and some minor elements (Al, Ca, etc.), were oxidized in the later stages. If oxygen was scarce, it sufficed only to oxidize iron into FeO, which then passed into the silicate phase. The remaining iron, and also nickel, copper and other metals, were not oxidized but became concentrated in the metal phase.

The proportion of these phases in each asteroid was determined essentially by the difference between the amounts of iron and oxygen which remained uncombined with other elements, such as sulfur (FeS), silicon, magnesium, hydrogen (as SiO_2, MgO, H_2O) etc.

In group I this difference was the greatest and so the maximum amount of the metal phase, iron-nickel occurs in it; in group V this difference was the least and the amount of metal is, therefore, the smallest. Besides this, the greater the amount of iron remaining in the metal phase, the smaller quantity of it passed into the silicate phase as FeO, and at the same time the smaller must be the relative concentration of nickel in the metal phase.

This, evidently, explains the relation among these components in each asteroid, which was named Prior's "primary" relation.

This explanation has certain features in common with that given by Prior for the relations in chondrites, but there is a certain basic difference. Prior thought this relation to be due to progressive oxidation of a single magma in a planet, while here the relationships have been considerably changed and broadened and are considered to be the results of a complex process of differentiation of matter first in the protoplanet stage and later in individual asteroids.

The data on the composition of asteroids suggest that besides the five asteroids from which the now known types of meteorites were formed, there may have existed other asteroids of intermediate composition.

Besides, it is possible to assume the existence of analogous bodies to the right of group V and to the left of group I (Figure 4). In the former, further increase in oxygen content can be expected and also a sharper increase in hydrogen with decrease in Fe, Mg and Si. Perhaps such bodies, which judging by their composition should have been farther from the sun, i. e., nearer to Jupiter, gave rise to comets?

On the other hand, bodies "before group I" should have had a relatively larger iron-nickel core with low nickel content. Fragments of such bodies falling upon the earth would have been iron meteorites with Ni content below 4%. In this connection, attention should be directed to finds of "native" iron not identified as meteoritic merely because of its low nickel content.

Such are a few general deductions based on the variation in the chemixal composition of meteorites.

In the next paper a discussion of the main characteristics of the mineral composition and structure of meteorites will be given which may confirm and extend the conclusions concerning the origin of meteorites.

The Meteorite Committee, Academy of Sciences, USSR

Received December 7, 1956

LITERATURE CITED

[1] G. T. Prior, Mineral. Mag. 18, 26 (1916).

[2] G. T. Prior, Mineral. Mag. 19, 51 (1920).

[3] L. J. Spencer, Mineral. Mag. 24, 296 (1936).

[4] P. N. Chirvinskii, Zap. Vserosissk. Mineral. Obshchestvo 68, 4, 580 (1939).

[5] E. P. Henderson and S. H. Perry, Smithsonian Inst. Publs. No. 3714, 235 (1943).

[6] H. Brown and C. Patterson, J. Geol. 55, 405; 508 (1947); 56, 85 (1948).

[7] J. M. Klotz, Science 109, 248 (1949).

[8] B. Iu. Levin, Meteoritika, 11, 47 (1954).

[9] H. C. Urey and H. Craig, Geochim. et Cosmochim. Acta, 4, No. 1/2, 36 (1953).

[10] A. A. Iavnel', Doklady Akad. Nauk SSSR, 102, 3, 477 (1955).

[11] A. A. Iavnel', Geokhimiia No. 2, 78 (1956).

[12] H. B. Wiik, Geochim. et Cosmochim. Acta, 9, No. 5/6, 279 (1956).

VARIATIONS OF THE MEAN LATITUDE OF POLTAVA, MIZUSAWA, RICHMOND AND WASHINGTON

E. I. Obrezkova

With the aim of discovering the causes of significant variations of mean latitude, parallel observations made on two instruments standing side by side at Poltava and Mizusawa are discussed, and in addition, the variations of the latitudes of Richmond and Washington, located at approximately the same meridian, are compared. The relation of these variations to the program of observations and the instrument is investigated. The conclusion is drawn that variations of mean latitude arising in the form of waves of an irregular character depend not on the program of observations but on certain defects in the instrument, which produce a systematic accumulation of observational errors. The latter bring about a fictitious variation of the mean latitude, which sometimes takes place abruptly. Since variations of mean latitude depend to a considerably extent on the instrument, they should be eliminated in calculating the coordinates of the instantaneous pole of the Earth, which unfortunately has not been done up to now by the Central Bureau of the International Latitude Service.

On the basis of an investigation of a large quantity of observational material by many authors, it is noted that the mean latitude varies. In particular, in [1] from a discussion of more than 140,000 latitude determinations we revealed a variation of the mean latitude of three international stations. As the results have shown, there is a progressive trend in the variation of latitude, and at the same time considerably larger variations are observed in the form of waves of an irregular character, both types of variation occurring in different ways at different stations, although the same stars are observed.

If the same observational material that was used in [1] is reduced on the proper motion system of Boss, then the mean latitude over a period of 34 years, for example, varies at each station by about 0".2. Therefore, the linear trend in the mean latitude variation depends partly on incorrect values of the stellar proper motions. This dependence of the variation of latitude was noted in the work [2] of Schlesinger, and recently attention was also called to it in the work [3] of Melchior. Because of inaccuracy of the stellar proper motions a fictitious secular variation of the mean latitudes arises.

As for the origin of the waves of irregular character referred to above, this is an extremely complex problem. The causes of these variations have not yet been explained. A. Ia. Orlov points out that these variations are of a local character and require special study [4]. However, the observations of the International Latitude Service do not permit a solution of this problem, since, in our opinion, it would be advisable for this purpose to conduct special observations on the general program at several locations on adjacent instruments. From what is stated below it would appear necessary to ascribe the waves of irregular character in the mean latitude variation to certain instrumental defects. This was especially well demonstrated by the observations at Mizusawa from 1943 to 1950.

In the present paper we report some results of our investigation of the dependence of latitude variation on the instrument and the observational program. For this purpose we made use of results of observations made at the following places:

1. At Poltava, where observations on the Zeiss and Bamberg zenith telescopes were carried out on the general program from 1949.8 to 1953.8.

2. At Richmond and Washington, where determinations of latitude were made on photographic zenith tubes from 1949.0 to 1953.0 on various programs.

3. Observations of latitude made at Mizusawa from 1942.8 to 1949.7 on a floating zenith telescope and a zenith telescope of the international type. The floating instrument was located 17 m west of the visual one on the same parallel 39° 08' 03".

§ 1. Variations of Latitude at Poltava from Observations on the Zeiss and Bamberg Zenith Telescopes

Observations at Poltava, as is well known, are made with two instruments, on the general program proposed by E. P. Fedorov and described in detail in his work [5]. In comparison to existing programs of latitude observations, this program is far better suited to the investigation of mean latitude variation, for which observations and their reduction require very much less labor than the reduction of observations of the International Latitude Service. In the Poltava program, principal attention was given to observation of two groups of Talcott pairs with mean right ascensions of 6^h and 18^h. One group, A, is observed every two hours for half the year, when it is replaced by the second group, C, which is observed every two hours for the second half of the year. Owing to the small effect of precession on the declinations of pairs of these groups, observations of them may extend over many decades. Besides, these groups comprise pairs of stars having very small differences in zenith distance, therefore the latitudes will be practically free of systematic errors arising from the inaccuracy in the assumed value of the screw constant.

The connection between the fundamental groups is established by mean of observations of two other groups, having mean right ascensions near 0^h and 12^h.

Observations of only the two groups A and C give a continous sequence of normal points for plotting curves of the latitude fluctuation. We plotted the latter for six years. The interval of time between mean dates of the normal points turned out to be 0.06 year, and each point represents, on the average, 48 pairs of observations. Besides, the observed latitudes were reduced to the centers of the corresponding groups A and C by means of corrections obtained from the general reduction of all observations by the chain method. From our curves of instantaneous latitude, plotted separately for the Zeiss and Bamberg instruments, the values of the latitude were taken at every 0.1 year, which wa shall call $\varphi_1, \varphi_2, \varphi_3, \varphi_4, \ldots, \varphi_n$. From the latter we initially took a mean of ten consecutive values, i. e., we calculated the values of the following expressions:

$$\frac{1}{10}\sum(\varphi_1+\varphi_2+\varphi_3+\cdots+\varphi_{10}),$$
$$\frac{1}{10}\sum(\varphi_2+\varphi_3+\varphi_4+\cdots+\varphi_{11}),$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

etc.

Each such mean is free of annual nonpolar latitude variation Z for the instant t + 4.5, where t is the initial instant of the time interval of the corresponding ten consecutive values of the latitude, t being expressed in tenths of a year. The results are listed in Table 1, where each number in Columns 2 and 4 contains

$$\Delta\varphi' = \frac{1}{10}\sum_{i}^{i+9}\varphi - 49^\circ 36' 12''.00$$

($\Delta\varphi'$ given in 0".01) for the two instruments, respectively, where i stands for the serial number of the latitude. We formed the differences $\Delta\varphi'$ (Column 2 minus Column 4) from observations on the two instruments and took the average of them. In Column 6, Table 1, are given the deviations of each difference from the mean. The numbers in Column 6 are free of secular and periodic variations of latitude, caused by motion of the pole. They are also free of the nonpolar periodic variation Z, because the latitudes, as was mentioned above, were taken in the mean for a year for each instrument.

Since the numbers of Column 6 were obtained from observations of the same stars on the two instruments, they do not depend on errors in declination. It is also impossible to explain their origin on the basis of the relative locations of the instruments (in the sense of differences in atmospheric effects, and effects of local peculiarities in the earth's crust) because the instruments stand side by side. It remains only to relate these quantities to those errors of observation which are not the same for both instruments. If not, i. e., it may be that the secular and periodic variations of latitude have not been entirely eliminated from these differences, then the variations of mean latitude calculated below from observations on each instrument will display this quantity in their corresponding differences. Using values of the instantaneous latitude for each tenth of a year, we calculated the secular variations of latitude by the method of A. Ia. Orlov [6]. For each instrument the values of the mean latitude are represented by the expression: $\psi = 49^\circ 36' 12''.00 + \Delta\psi$. The values of $\Delta\psi$ obtained (in $0''.01$) are listed in Table 1 (see Columns 3 and 5), and are also represented graphically in Figure 1 for each instrument separately. The differences of the mean latitudes of these instruments, after eliminating the constant component, are given in Column 7 of the same table.

Comparison of the numbers in Columns 6 and 7, obtained by two methods, shows that they have the same trend. This implies that the magnitudes obtained in Column 6 are free from secular as well as periodic variations of latitude. On the other hand, this implies that A. Ia. Orlov's method of deriving secular latitude variations eliminates the periodic latitude variations, both polar and nonpolar.

The numbers of Column 6, and therefore, of Column 7 as well, depend, in the first place, on random observational errors of both instruments and those adjustments and smoothings made in the calculation and plotting of the curves, in the second place, on certain systematic variations originating in the instruments themselves.

The mean square of the residuals of the latitude differences, obtained by the first method (Column 6) equals $\pm 0''.024$, and by the second method (Column 7) it equals $\pm 0''.021$. These numbers indicate that the method of A. Ia. Orlov slightly smoothes the variation of the differences and hence the latitude variations themselves.

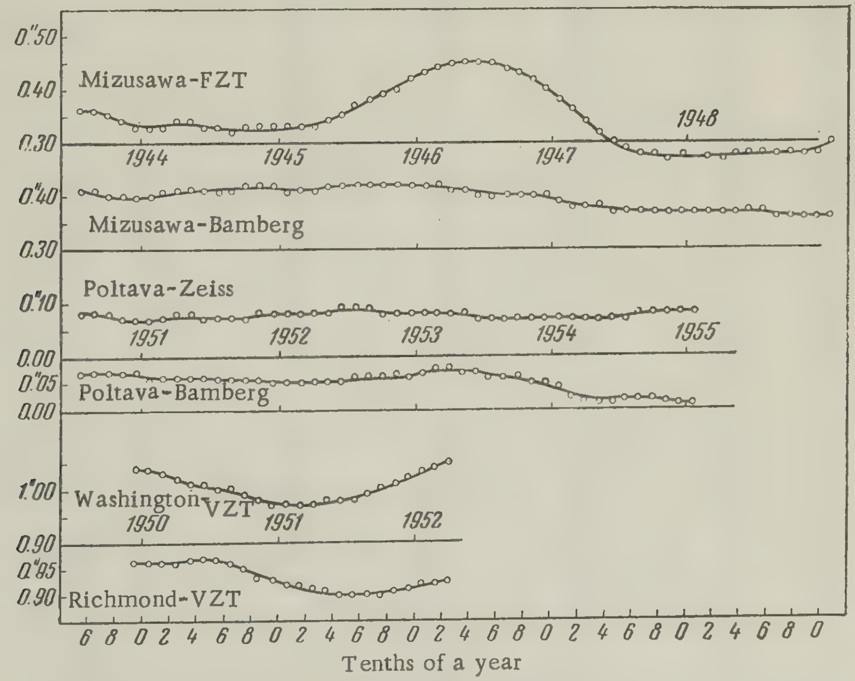


Fig. 1. Variations of mean latitude from observations with visual and photographic instruments.

TABLE 1

Differences in the Values of the Mean Latitude (0″.01) Obtained by Two Methods from Observations on the Two Zenith Telescopes at Poltava (Constant Component Eliminated)

Date	Zeiss		Bamberg		Residuals	
	Δφ′	Δψ	Δφ′	Δψ		
1	2	3	4	5	6	7
1950.25	108	—	108	—	—	—
35	110	—	110	—	—	—
45	113	—	112	—	—	—
55	113	108	112	107	+2	+2
65	112	108	110	107	+1	+2
75	110	108	108	107	+1	+2
85	108	108	106	107	+1	+2
95	105	108	104	107	+2	+2
1951.05	104	108	102	106	+1	+1
15	104	108	102	107	+1	+2
25	105	109	103	107	+1	+1
35	106	109	105	107	+2	+1
45	109	108	106	107	0	+2
55	113	108	110	106	0	+1
65	114	108	111	106	0	+1
75	114	108	111	106	0	+1
85	113	109	109	106	—1	0
95	111	110	107	105	—1	—2
1952.05	108	110	104	105	—1	—2
15	106	109	101	105	—2	—1
25	105	109	99	105	—3	—1
35	105	109	100	105	—2	—1
45	106	109	101	105	—2	—1
55	107	109	103	106	—1	0
65	110	109	107	106	0	0
75	112	108	110	106	+1	+1
85	113	108	112	106	+2	+1
95	113	108	112	106	+2	+1
1953.05	111	108	110	107	+2	+2
15	109	108	108	108	+2	+3
25	106	108	106	108	+3	+3
35	104	108	104	107	+3	+2
45	103	107	104	107	+4	+3
55	103	107	104	106	+4	+2
65	104	107	104	106	+3	+2
75	106	107	105	106	+2	+2
85	108	107	106	105	+1	+1
95	110	107	107	105	0	+1
1954.05	111	107	108	104	0	0
15	111	107	107	102	—1	—2
25	109	107	105	101	—1	—3
35	107	107	102	101	—2	—3
45	106	107	98	101	—5	—3
55	104	107	96	102	—5	—2
65	103	108	94	102	—6	—3
75	104	108	96	102	—5	—3
85	106	008	99	101	—4	—4
95	108	108	102	101	—3	—4
1945.05	110	108	105	101	—2	—4
15	112	—	107	—	—	—
25	113	—	107	—	—	—

So, from comparison of the numbers of Columns 6 and 7 let us note that, in the first place, the differences of instantaneous latitudes, being, as has been shown, differences of observational errors, remain almost entirely in the differences of the mean latitudes and therefore fictitiously vary the values of the mean latitude; in the second place, A. Ia. Orlov's method of deriving the secular latitude variations to some extent reduces these errors; in the third place, the differences in the values of the mean latitude obtained by A. Ia. Orlov's formula from observations on two instruments may be considered differences of instantaneous latitudes, free from the quantities Z of those instruments. Thus, comparison of mean latitude variations for two neighboring instruments may show an advantage of some instruments over others. In order to satisfy ourselves further of this, we continue in Sections 2 and 3 a similar investigation of the results of observations carried out at the other points already mentioned earlier.

§ 2. Variations of Latitude of the Mizusawa Station from Observations in the Period 1942.8 to 1949.7.

In the derivation of the secular variations of latitude of Mizusawa, the raw material consisted of the results of latitude observations made on a floating photographic zenith telescope (FZT) and a zenith telescope of the international type (VZT), located in Mizusawa on the same parallel, at a distance of 17 m from each other. Parallel observations on the two instruments were begun in 1940 on the international program. But several stars turned out to be much too faint for the FZT. Then in September 1942 a new program was introduced for this instrument, also consisting of 12 groups of 8 star pairs each. The total number of pairs in the old program remained 32. All stars of the floating photographic zenith telescope program were taken from the Boss catalog.

The results of the observations made on the instruments referred to were published in [7] in the form of group-mean values of the instantaneous latitudes referred to the mean date, which is the same for both instruments, the interval of time between mean dates of observation being 30 days. Each group was obtained as a mean of 103 pairs of observations for one instrument, the FZT, and 107 for the second, the VZT. On the basis of the grouped values of the latitudes, we obtained graphically for each tenth of a year the values of the instantaneous latitude. From the latter we took the average values for the year and thereby freed them from the nonpolar variations Z. The results are given in Table 2 for each instrument separately (Columns 2 and 4). The deviations of the latitude differences from the mean (Column 2 minus Column 4) are presented in Column 6 of the same Table 2. It is impossible to relate the latter to the real variations of latitude since both instruments stand side by side.

Using the formula of A. Ia. Orlov, we calculated the variations of mean latitude ψ from observations on each instrument and formed the differences of the mean latitudes. The results are presented in Table 2, in which Columns 3 and 5 contain the values of ψ respectively for the two instruments, and the differences of these ψ are entered in Column 7. The variations of mean latitude for each instrument separately are shown in the graph (Figure 1, the two upper curves).

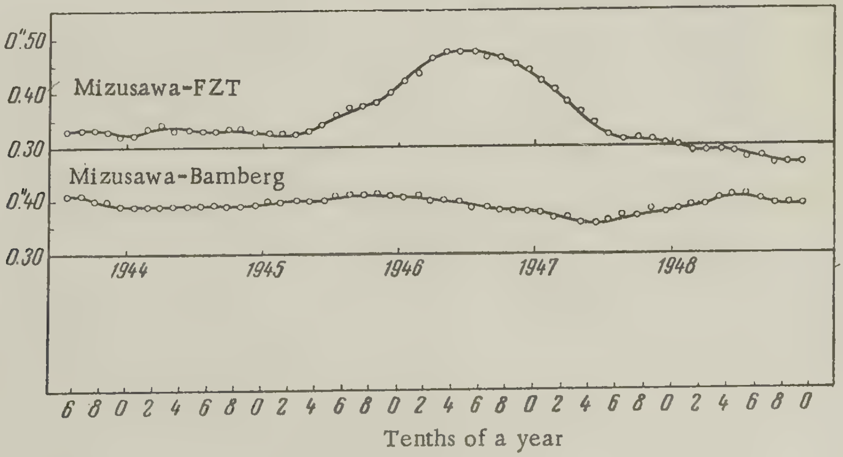


Fig. 2. Variations of mean latitude from observations of a total of 32 pairs on two instruments.

TABLE 2

Differences in Mean Latitude (in 0″.01) Obtained by Two Methods from Observations on the Floating Photographic and Visual Telescopes at Mizusawa (Constant Component Eliminated)

Date	FZT		VZT		Residuals	
	φ′	ψ	φ′	ψ		
1	2	3	4	5	6	7
1943.25	3″36	—	3″41	—	—	—
35	36	—	41	—	—	—
45	36	—	42	—	—	—
55	37	3″36	42	3″41	0	0
65	37	36	42	41	0	0
75	37	35	42	40	0	0
85	36	34	42	40	+1	+1
95	35	33	41	40	+1	+2
1944.05	34	33	39	40	0	+2
15	32	33	38	41	+1	+3
25	31	34	38	41	+2	+2
35	30	34	38	41	+3	+2
45	31	33	40	41	+4	+3
55	32	33	41	41	+4	+3
65	34	32	42	41	+3	+4
75	35	33	44	42	+4	+4
85	35	33	44	42	+4	+4
95	35	33	44	42	+4	+4
1945.05	35	33	44	41	+4	+3
15	34	33	43	41	+4	+3
25	33	33	41	41	+3	+3
35	32	34	39	42	+2	+3
45	32	35	38	42	+1	+2
55	32	37	39	42	+2	0
65	34	38	40	42	+1	—1
75	37	39	41	42	—1	—2
85	41	40	43	42	—3	—3
95	44	42	44	42	—5	—5
1946.05	46	43	45	42	—6	—6
15	48	44	45	42	—8	—7
25	50	45	44	41	—11	—9
35	49	45	42	41	—12	—9
45	48	45	40	40	—13	—10
55	46	45	38	40	—13	—10
65	45	44	37	40	—13	—9
75	43	43	36	40	—12	—8
85	41	42	36	40	—10	—7
95	40	40	38	40	—7	—5
1947.05	39	38	40	39	—4	—4
15	38	36	41	38	—2	—3
25	36	34	42	38	+1	—1
35	33	32	41	38	+3	+1
45	31	30	40	37	+4	+2
55	28	29	38	37	+5	+3
65	25	28	35	37	+5	+4
75	23	28	33	37	+5	+4
85	22	27	32	37	+5	+5
95	22	28	32	37	+5	+4
1948.05	24	27	34	37	+5	+5
15	27	27	36	37	+4	+5
25	30	27	38	37	+3	+5
35	32	28	40	37	+3	+4
45	32	28	42	37	+5	+4
55	32	28	42	37	+5	+4
65	31	28	40	36	+4	+3
75	29	28	38	36	+4	+3
85	26	28	35	36	+4	+3
95	24	28	32	36	+3	+3
1949.05	23	30	30	36	+2	+1
15	23	—	30	—	—	—
25	26	—	32	—	—	—
35	30	—	35	—	—	—

A comparison of the latitude differences obtained by the two methods (Columns 6 and 7, Table 2), indicates that their trend is the same. In this way the differences obtained from the mean latitudes, computed by the method of A. Ia, Orlov, are found to be smoothed more in comparison to the differences of instantaneous latitude, free from Z; the mean square of the residuals of the latitude differences obtained by the first method (Column 6) is ± 0″.055, and by the second (Column 7) is ± 0″.045.

The quantities listed in Columns 6 and 7 do not depend on differences in the observational programs. We were convinced of this by analyzing the results of observation of the total of 32 pairs on the instruments mentioned. The values of the latitudes obtained from these pairs are given in [8]. We reduced them to the mean system of the deviations and proper motions of the whole program, making use of the corrections $\Delta\delta$ and $\Delta\mu'$ taken from [9]. After this we formed the normal values of the latitudes and the corresponding mean dates of the observations for each instrument separately. Each normal point represents, on the average, 49 latitudes obtained by means of the visual zenith telescope and 38 by means of the floating one. The intervals of time between mean dates were the same for each instrument and equal to 32 days. Having the normal latitudes for the mean dates of observation, we graphically determined their values at each tenth of a year for each instrument, after which we calculated the mean latitude by the method of A. Ia. Orlov. The results are listed in Table 3 and represented graphically in Figure 2. Comparison of the latter with Figure 1 shows that the mean latitude variations do not depend markedly on the program of observations. These variations arise principally from some defects in the instruments which cause errors in the observations and, as the results obtained indicate, the floating zenith telescope is less reliable in this respect. Observations made with it apparently are greatly subject to systematic errors which markedly and abruptly show an effect on the variation of mean latitude. Owing to their influence, as is indicated by the magnitudes of Column 3 for the FZT in Table 2, the variation of latitude in the period 1946.4 to 1947.8 reaches 0″.18. Since this variation of latitude for Mizusawa corresponds to about 6 m on the earth's surface, it is quite obvious that, with a distance of 17 m between the instruments, it is impossible to explain this magnitude by variation of the relative positions of the instruments. It is also impossible to explain it by different programs of observations, since, having separated out all 32 pairs from the two different programs we obtained the same variations of latitude.

TABLE 3

Variations of Mean Latitude at Mizusawa from Observations of 32 Pairs on Two Instruments

Date	FZT	VZT	Date	FZT	VZT
	ψ	ψ		ψ	ψ
1943.55	3″33	3″41	1946.35	3″47	3″40
65	33	41	45	47	40
75	33	40	55	47	39
85	33	40	65	46	39
95	32	39	75	46	38
1944.05	32	39	85	45	38
15	33	39	95	44	38
25	34	39	1947.05	42	38
35	33	39	15	40	37
45	33	39	25	38	37
55	33	39	35	36	36
65	33	39	45	34	36
75	33	39	55	32	36
85	33	39	65	31	37
95	33	39	75	31	37
1945.05	32	40	85	31	38
15	33	40	95	30	38
25	36	40	1948.05	30	38
35	33	40	15	29	39
45	34	40	25	29	39
55	36	41	35	29	40
65	37	41	45	29	41
75	37	41	55	28	41
85	38	41	65	28	40
95	40	41	75	27	39
1946.05	42	41	85	27	39
15	43	41	95	27	39
25	46	40			

Thus the data of Tables 2 and 3, and also the graphs of Figures 1 and 2, clearly indicate that at Mizusawa, for the same interval of time, the observations on the floating zenith telescope give variations of mean latitude entirely different in character and magnitude from those yielded by observations on the visual telescope. Thus, in the interval from 1945.3 to 1946.3, the variations of mean latitude from observations on the floating zenith telescope attain the magnitude 0".11, while observations on the second instrument indicate that the latitude remained almost unchanged. It is not possible to explain this difference in variations of latitude, as was noted above, either by variation in relative location of the instruments, or by a difference in programs of observation for the two instruments. The origin of such abrupt changes in latitude (0".18) must be sought in defects of the floating zenith telescope.

§ 3. The Mean Latitude at Richmond and Washington from Observations in the Period 1949.0 to 1953.0.

Observations at Richmond and Washington were carried out on photographic zenith tubes on different programs and were published in [10] and [11] in the form of normal values of the instantaneous latitudes for each instruments. These values of the latitudes were given for each 0.05 year and were reduced to the mean epochs of the observations, the same for the two instruments. We found the values of the instantaneous latitudes for each tenth of a year. After this, Table 4 was compiled similarly to Tables 1 and 2 of the preceding paragraphs.

TABLE 4

Differences of Mean Latitude (in 0".01) Obtained by Two Methods, from Observations on Zenith Tubes at Washington and Richmond (Constant Component Eliminated)

Date	Washington		Richmond		Residuals	
	$\Delta\varphi$	$\Delta\psi$	$\Delta\varphi$	$\Delta\psi$		
1949.65	102	—	91	—	—	—
75	103	—	99	—	—	—
85	105	—	96	—	—	—
95	107	104	99	97	—1	0
1950.05	108	104	102	97	+1	0
15	108	103	104	97	+3	+1
25	107	102	103	96	+3	+1
35	105	101	100	97	+2	+3
45	101	101	97	97	+3	+3
55	98	100	94	97	+3	+4
65	95	100	92	96	+4	+3
75	93	99	90	95	+4	+3
85	93	98	90	93	+4	+2
95	95	97	91	93	+3	+3
1951.05	97	97	92	92	+2	+2
15	99	97	93	92	+1	+2
25	100	97	94	91	+1	+1
35	101	98	94	91	0	0
45	101	98	93	90	—1	—1
55	100	98	92	90	—1	—1
65	98	99	89	90	—2	—2
75	97	100	87	90	—3	—3
85	97	101	86	91	—4	—3
95	97	102	86	91	—4	—4
1952.05	99	103	87	92	—5	—4
15	102	104	90	92	—5	—5
25	106	105	93	92	—6	—6
35	109	—	96	—	—	—
45	111	—	98	—	—	—
55	113	—	99	—	—	—

Table 4 contains in Columns 2 and 4 the variation of latitude $\Delta\varphi = \frac{1}{10}\sum_{i}^{i+9}\varphi - 38^\circ55'16''.00$ for Washington and $\Delta\varphi = \frac{1}{10}\sum_{i}^{i+9}\varphi - 25^\circ37'27''.00$ for Richmond, where φ are the values of the instantaneous latitude, and i is its serial number.

In Columns 3 and 5 are listed variations of mean latitude $\Delta\psi = \psi - 38^\circ55'16''.00$ and $\Delta\psi = \psi - 25^\circ37'27''.00$ respectively for Washington and Richmond. The values of the mean latitude, calculated by A. Ia. Orlov's method from observations at these points, are represented also in the graph (Fig. 1). The differences of mean latitudes (Washington minus Richmond) obtained by the first method (the numbers of Column 6) have the same trend as the differences of latitudes obtained by the second method (numbers of Column 7). This result is similar to that obtained from observations on the two neighboring instruments at Poltava, and also at Mizusawa (see Sections 1 and 2). However, the observations at Poltava and Mizusawa showed that these differences depend on certain instrumental defects which give rise to a fictitious variation of mean latitude. The differences of the latitudes of Richmond and Washington were obtained from observations carried out on two instruments of the same type. The data of Table 4 and the graph (Figure 1, two lower curves) indicate the same character of mean latitude variation at Richmond and Washington.

A. Ia. Orlov remarks in his work that the variation of mean latitude ψ does not depend on motion of the pole. In this connection he gives the example of the variation of mean latitude at two stations — Greenwich (Gr) and Carloforte (Ca). Both stations, like Richmond and Washington, lie on nearly the same meridian on one side of the pole. For these stations A. Ia. Orlov found the quantities which we quote in the following Table 5 (ψ expressed in $0''.01$).

TABLE 5

Date	Gr	Ca	$\psi_{Gr} - \psi_{Ca}$	Date	Gr	Ca	$\psi_{Gr} - \psi_{Ca}$
	ψ	ψ			ψ	ψ	
1917.0	—1	—7	+6	1922.5	0	—	—
.5	0	—7	+7	1923.0	—2	—	—
18.0	0	—7	+7	.5	—4	—13	—9
.5	0	—5	+5	24.0	—5	—13	+8
19.0	+1	—3	+4	.5	—4	—14	+10
.5	+1	—2	+3	25.0	—3	—14	+11
20.0	+2	—1	+3	.5	—1	—13	+12
.5	+4	+2	+2	26.0	—1	—14	+13
21.0	+2	+2	0	.5	+1	—10	+16
.5	+1	+2	—1	27.0	+4	—14	+18
22.0	+1	—	—	.5	+3	—13	+16

"From these numbers it is evident that the mean latitude variations at Greenwich and Carloforte are entirely different. From 1917.0 to 1921.5 at Carloforte ψ increases all the time, while at Greenwich for the same period this quantity having reached maximum in 1920.5, already markedly diminished in a year. From 1924 at Greenwich the mean latitude, all the while increasing, became larger by $0''.09$ in three years, while at Carloforte during the same period it remained almost constant" ([12], p. 7).

In Mizusawa the same type of instruments were set up as in Greenwich and Carloforte. As was noted above, the variation of mean latitude at Mizusawa on each of the two instruments (set up side by side) indicated that the observations on the floating zenith telescope give variations of latitude entirely different in character from those yielded by observations on the visual one. A similar phenomenon is observed in the latitude variations listed in Table 5. The disagreement of the curves of mean latitude variation at Greenwich and at Carloforte may, in all likelihood, be explained by the fact that observations on the floating zenith telescope (at Greenwich) are burdened by larger errors than are observations on the visual zenith telescope (at Carloforte). Therefore observations on the floating zenith telescope will not allow us to establish the dependence of ψ on motion of the pole if certain defects in the instrument, producing errors in the observations, are not eliminated.

Turning to a study of the results of observations at Washington and Richmond, let us note that the similarity of the curves of mean latitude variation (Figure 1) is indicative of a dependence of ψ on motion of the pole. However, this result may not yield conclusive evidence of such a relation, since the observations at these points cover a short interval of time.

As mentioned above, A. Ia. Orlov write that the causes of the irregular variations of mean latitude have a local character. As the present work has shown, the cause of these variations is concealed not in some peculiarities of some change in position of the land, but in some peculiarities of the instrument. The graphs (Figures 1 and 2), plotted from the data of Tables 1, 2, 3 and 4, give evidence of the fact that the variations of mean latitude depend on the instrument; observations on visual instruments, the Zeiss and Bamberg zenith telescopes at Poltava, give variations of the mean latitude represented graphically by a nearly straight line, observations on the Bamberg at Mizusawa indicate that the mean latitude from 1943.6 to 1949.0 remained constant, while observations on the floating zenith telescope adjacent to the Bamberg during the same period give variations of the mean latitude in the form of waves of an irregular character. Observations at Washington and Richmond on photographic zenith tubes also give variations of the mean latitude in the form of abruptly changing curves. It is entirely obvious that photographic zenith tubes and the floating zenith telescope give rise to greater variations of the mean latitude than do visual instruments.

Since the mean latitude variations depend to a considerable extent on the instrument, as the present work has shown, secular variations of the mean latitudes of stations should be eliminated in calculating the coordinates of the pole, as was suggested for the first time by A. Ia. Orlov [6] and acknowledged by Soviet astronomers, but the Central Bureau of the International Latitude Service to the present time calculates the coordinates of the instantaneous pole without eliminating the secular variations in latitude of the stations.

The results obtained emphasize the necessity of studying instrumental errors. In our opinion, only simultaneous study of the data of parallel observations on two instruments will make possible the discovery of the cause of significant mean latitude variations. Therefore the organization of such observations, and hence also their continuation where they are already being made, is highly desirable. The obtaining of observational data at these points over a longer interval of time will allow us to count on progress in the solution of this complex problem — the explanation of the causes of variation in mean latitude.

Poltava Gravimetric Observatory,
Academy of Sciences Ukraine SSR

Received December 24, 1956

LITERATURE CITED

[1] E. I. Obreskova, Astron. Tsirkuliar AN SSSR, No. 143 (1953).

[2] F. Schlesinger, Astron. J. 34, No. 798 (1922).

[3] P. Melchior, Communic, de l'Obs. Royal de Belgique, No. 79 (1955).

[4] A. Ia. Orlov, Doklady Akad. Nauk SSSR 28, No. 4 (1951).

[5] Trudy Poltavskoi Gravimetricheskoi Observatorii 4 (Kiev, 1951).

[6] A. Orloff, Harmonic Analysis of the Latitude Observations (Odessa, 1925).

[7] T. Hattori, Draft Reports, 3 (Cambridge, 1952).

[8] T. Hattori, Publ. Internat. Latitude Observatory of Mizusawa, 1, No. 1 (1951).

[9] T. Hattori, Publ. Internat. Latitude Observatory of Mizusawa, 1, No. 2 (1953).

[10] P. Sollenberger, Astron. J. 57, 122 (1952).

[11] W. Markowitz, Astron. J. 59, 132 (1954).

[12] A. Ia. Orlov, Biulleten' GAISh No. 8 (1941).

THE FUNDAMENTAL GRAVIMETRIC STATION AT THE P. K. SHTERNBERG STATE ASTRONOMICAL INSTITUTE (GAISh)

N. P. Grushinskii

This article describes the determination of the fundamental gravimetric station at the new building of the Astronomical Institute located in the Lenin Mountains. It also gives the value of gravity at this station, and contains information on the connection between the original station at the Moscow Aerogeodetic Organization and the main initial gravimetric station at the Shternberg State Astronomical Institute.

There are four main auxiliary gravimetric stations in the USSR; these are directly related to the Postdam station, where, between 1898 and 1904, Kuenen and Furchtwangler carried out the determinations of the absolute force of gravity on which the entire world's gravimetric survey is based. These four stations are:

TABLE 1

Name of Station	Years of Correlation	H, m	Value
Moscow, GAISh	1913	145	981559.0 ± 0.74
Leningrad, Pulkovo	1897, 1901, 1930	71	981900.5 ± 0.54
Kazan'	1908, 1913	76	981558.7 ± 0.68
Poltava, Gravimetric Observatory	1927	146	981006.4 ± 0.67

Values of the force of gravity shown in this table are interrelated with one another and with Potsdam (N. N. Pariiskii, Acceleration of Gravity at the Main Auxiliary Stations of the USSR; Pulkovo, Moscow, Poltava and Kazan'. Reports of the VTOGIRG, V. 1, 1935).

Through these stations the entire gravimetric survey of the Union is related to the Potsdam System of gravity determinations. Most of the Russian determinations were carried out in connection with the Moscow main auxiliary station. This station is located on the grounds of the Astronomical Observatory at Krasnaia Presnia (formerly at No. 5, Plavlik Morozov). A direct correlation exists between Potsdam and the eastern pillar of the auditorium. This correlation was established by Haaseman in 1913. Since this pillar was found to be unsuited for conducting observations, another pillar, located in the gravimetric (central) basement, was taken for the base station. The top area of this pillar lies at 2.9 m below the top area of the pillar in the auditorium, which caused a reduction in the force of gravity between the pillar in the auditorium and the pillar in the basement by + 0.7 mgl. (The reduction in the atmosphere minus the action of the intervening column of air). The value of the force of gravity for the lower pillar is shown in Table 1. A more exact determination of the difference between the forces of gravity for the two pillars was carried out by us in March, 1956. This difference was found to be equal to

0.5 mgl ± 0.04. In this work, five determinations were made with the use of a GAK-3 M gravimeter. Most of the pendulum gravimetric stations of the USSR were established in relation to this pillar.

For a long time a pillar in the western basement was also used as the base station. At this time this station has been lost; however, since it was located at the same level as the pillar of the gravimetric basement and only at a distance of a few meters to the southwest of it, it can be assumed that the difference in the force of gravity between the two pillars lies within the limits of instrumental accuracy.

After 1940 the auxiliary gravimetric station at the Moscow Aerogeodetic Organization located in the MAGP basement (6 Shelaputinskii Lane) became the basic gravimetric station for the pendulum surveys. This station was interrelated by a pendulum apparatus with the basic auxiliary station GAISh. In 1951 their interrelationship was reestablished by means of gravimeters. This work was done by the author and the scientific member of the TsNIIGAiK,[*] L. A. Govorova. The difference between the main auxiliary station GAISh and the main basic station MAGP, considered in the form of $g_{GAISh} - g_{MAGP}$, is shown in Table 2.

TABLE 2

Date January, 1952	Norgard Gravimeter						
	№ 271	№ 327	№ 1355	№ 1137	№ 1080	№ 1357	mean in mgl
14	8.7±0.25	9.0±0.04	8.9±0.17				8.9
16	—	—	—	8.8±0.06	9.1±0.10	9.0±0.10	9.0
17	9.2±0.09	—	9.0±0.06	—	—	—	9.2
Mean in mgl	9.0	9.0	9.0	8.8	9.1	9.0	9.0±0.04

Each value of the difference in the force of gravity shown here represents a mean of three determinations completed with a given apparatus at a given time. The evaluation was made by comparing the deviations of the differences obtained in each determination with the mean obtained with a given instrument at a given time. The evaluation of the final result was made by comparing the deviations of the mean values of the difference found with all the instruments used at a given time with the general mean.

In this way the force of gravity at the main basic station MAGP was found to be equal to 981559.0 ± 0.74 — — 9.0 ± 0.04 = 981550.0 ± 0.75 mgl.

At this time, owing to the relocation of the Astronomical Institute in the Lenin Mountains, a new fundamental gravimetric station was established. To this end, the value of the force of gravity had to be transferred from the pillar in the gravimetric basement of the Observatory in Presnia to the GAISh Gravimetric Laboratory in Lenin Mountains. This effort was started in the spring of 1955 by the co-workers of the Department of Gravimetry of the Mechanical-Mathematical College, I. G. Koreneva, V. L. Panteleev and I. N. Kantsova, under the direction of the author. The observations were conducted at the round foremost left pillar of the Gravimetric Laboratory at the Observatory in the Lenin Mountains. The plan showing the location of this pillar is presented on Figure 1.

The transferring of the force of gravity from the GAISh gravimetric basement on Presnia to the GAISh Gravimetric Laboratory in the Lenin Mountains was accomplished with the help of four gravimeters. Table 3 shows the types of the gravimeters used, the number of days spent on observations, and the number of times each gravimeter was used per day.

The instruments used were so adjusted as to enable us to make allowances for the displacement of the zero-point when the instruments were used independently. The selection of proper instruments was carried out by making three observations with each instrument: at the first station, at the second, and then again at the first; or at the second station, at the first and again at the second station. In this work, all the instruments were transported on a GAZ-51[**] truck. The observations were carried out simultaneously with all the gravimeters. Several observations were taken consecutively at each station during every day. The time interval between the observations at different points ranged from one to one-and-a-half hours. At the time of the observations the outside temperature and the temperature in the buildings in which the work was done were equal, so that the temperature of the apparatus varied very little. Table 4 shows the change in the temperature of the apparatus during the work.

---

[*] TsNIIGAiK = Central Scientific Research Institute of Geodesy, Aerial Surveying and Cartography.

[**] GAZ = Gor'kii Motor Vehicle Plant.

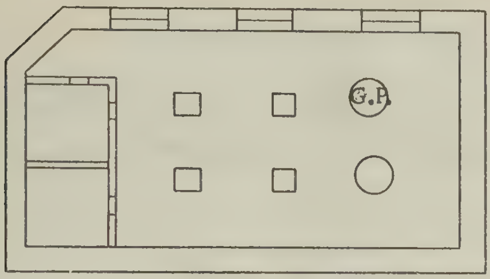


Fig. 1.

In working out the experimental results, corrections were used for the temperature, the zero-point displacement and the scale coefficient. The magnitude of hourly displacement of the zero-point was determined as the mean weight change for the zero-point during all observations during the day. In computing the final values of the difference in the force of gravity, new scale coefficients obtained during the calibration of the gravimeters were taken into consideration. The calibration of the Norgard and the CN-3 gravimeters was accomplished in two ways: by altering the inclination and by changing the turning angles with the help of the protractor. Moreover, Norgard gravimeter No. 1416 was calibrated for the Poltava and Moscow stations. The GAK-3M gravimeter was calibrated by means of changing its inclination.

TABLE 3

Gravimeter Type and number	Date June, 1955	No. of determinations
SN-3	9	2
No. 5135	11	3
	13	3
GAK-3M	9	2
No. 22271	11	3
	13	3
Norgard	9	2
No. 1416	11	3
	13	3
Norgard	9	2
No. 1418	11	3
	13	3

TABLE 4

Gravimeter	Date June	Time interval	Temperature changes
Norgard	9	5 hrs. 40 min	+ 2.2°
No. 1416	11	8 hrs.	+ 0.3
	13	6 hrs. 50 min.	+ 1.2
Norgard	9	5 hrs. 30 min.	+ 2.5
No. 1418	11	7 hrs. 30 min.	− 1.6
GAK-3M	9	6 hrs. 15 min.	+ 1.2
No. 22271	11	6 hrs. 40 min.	− 1.0
	13	6 hrs. 40 min.	+ 0.4

Plus sign corresponds to the rise of temperature, minus sign corresponds to the lowering.

The results of calibration are shown in Table 5. It can be seen from the table that the calibration of the quartz gravimeters with a twisting thread can be accomplished satisfactorily by all the methods, but that the results obtained by various methods exhibit systematic differences. We gave preference to the method which involved changing the inclination because the result obtained by this method was very near to the result of calibration for the stations with the known value of the force of gravity.

The accepted values of the scale coefficient are shown in the last column of Table 5. In Table 6 we show the measured differences in the force of gravity between the pillar at the gravimetric basement at Presnia and the pillar at the Gravimetric Laboratory in the Lenin Mountains, this difference being expressed in the form of $g_{\text{Presnia}} - g_{\text{Lenin Mts.}}$

Each difference of the force of gravity obtained with every instrument and entered in Table 6 shows a mean of the results obtained with a given instrument on a given day. The mean quadratic errors in these values show the errors in the mean result caused by the imperfect correlation of observations throughout a given day. We took the value of 39.5 mgl ± 0.08, obtained as a mean for the difference Δg by all the instruments, as the final result for the difference in the force of gravity between the gravimetric station at Presnia and the Lenin Mountains.

TABLE 5

Instrument	Date	Method	Scale coefficient	Accepted values
Norgard No. 1416	22.7.55	Inclination	0.9978	0.9984
	23.7.55	"	0.9976	
	30.5.55	Change of angles	0.9953	
	9.9.55	"	0.9952	
Norgard No. 1418	6—11.7.53	For Poltava-Moscow	0.9990	
	25.7.55	Inclination	0.9979	0.9981
	26.7.55	"	0.9983	
	4.6.55	Change of angles	0.9966	
	5.8.55	"	0.9948	
SN-3 No. 5135	2.8.55	Inclination	0.9993	
	6.6.55	Change of angles	0.9982	0.9988
GAK-3M No. 22271	14.3.55	Inclination	4.41	
	22.3.55	"	4.34	
	29.3.55	"	4.36	4.38
	20.4.55	"	4.43	

TABLE 6

Date	Values of Δg and of mean quadratic error in gravimetric determinations, in mgl				
	Norgard No. 1416	Norgard No. 1418	SN-3 No. 5135	GAK-3M No. 22271	daily mean
9.6.55	+39.4±0.10	+39.7±0.17	+39.3±0.14	+39.0±0.30	39.3
11.6.55	40.1±0.11	39.3±0.24	39.9±0.07	39.5±0.13	39.7
13.6.55	39.5±0.15	39.4±0.16	39.3±0.25	39.4±0.08	39.4
Mean for the instrument	39.7±0.19	39.5±0.16	39.5±0.22	39.3±0.16 mean39.5±0.08	

In evaluating the accuracy we accepted the mean quadratic error in the mean result, shown in the final line of Table 6.

Furthermore, we correlated the pillar at the Gravimetric Laboratory with the central pillar of the gravimetric basements. This correlating process was carried out on March 6, 1956, by four applications of gravimeter GAK-3M. The result of this operation showed the difference between the force of gravity at the gravimetric pillars of the Laboratory and the basement to be $g_{lab} - g_{basement} = -0.98 \pm 0.05$ mgl.

The value of the force of gravity in the Gravimetric Laboratory in the Lenin Mountains, as related to the Potsdam system, is equal to 981559.0 (± 0.74) − 39.5 (± 0.08) = 981519.5 (± 0.75) mgl.

The value of the force of gravity in the gravimetric basement in the Lenin Mountains, as related to the Potsdam system, is equal to 981559.0 − 39.5 + 1.0 = 981520 ± 10.75 mgl.

P. K. Shternberg State Astronomical Institute

Received June 19, 1956

THE INFLUENCE OF THE POSITION ERROR OF A SHIP ON THE ASTRONOMICALLY DETERMINABLE COMPASS CORRECTION

A. B. Marinbakh

The influence of the errors of the dead reckoning position of a ship on the compass correction, determined from the bearing of a celestial body, is discussed. Conditions under which the deviation of a ship from its course, caused by these errors, is negligible, are determined.

Compass correction (for either a magnetic or a gyroscopic instrument) is determined as the difference between the azimuth of a heavenly body (A) and its compass bearing (KП):

$$\Delta \text{К} = \text{A} - \text{КП}.$$

If computed from the coordinates of a dead reckoning position, this correction will be in error (see [3], pp. 78 and 485) by the magnitude*

$$\Delta(\Delta \text{К}) = \Delta\lambda(\sin\varphi - \cos\varphi\cos A\,\mathrm{tg}\,h) + \Delta\varphi\sin A\,\mathrm{tg}\,h,$$

where $\Delta\varphi$ and $\Delta\lambda$ represent the errors in the computed latitude and longitude, and $\underline{h}$ stands for the elevation of the heavenly body; azimuths and bearings are reckoned from the north to the east, and the longitude to the east of the prime meridian. The compass course (KK) produced by the error will be

$$\text{KK} = \text{IK} - \Delta\text{K},$$

and will be used in navigating the ship instead of the desired actual course (IK), either computed or taken from the map from the point designated by the coordinates of the ship to the point of its destination ([2] chapter 22); a ship following this course will deviate from the desired course by the angle $\Delta(\Delta\text{K})$.

By introducing the designations $\Delta w = \Delta\lambda\cos\varphi$ and $\Delta n = \Delta w\cos A - \Delta\varphi\sin A$, we transform the Expression (1) into:

$$\Delta(\Delta \text{К}) = \Delta w\,\mathrm{tg}\,\varphi - \Delta n\,\mathrm{tg}\,h. \qquad (2)$$

Because of the convergence of the meridians ($\Delta\lambda\sin\varphi$) with the increase of the latitude of a point, the influence of this error (Δw) on the error in compass correction (product $\Delta w\tan\varphi$ in [2]) increases rapidly. At large distances from the Equator the expression $\Delta w\tan\varphi$ cannot be neglected. Let us assume that the error due to the deflection (Δw) reached 30 miles, which is possible ([2], p. 326) after prolonged navigation without taking observations, after a storm, etc.; then the error in the compass correction will, under the best of conditions (at $h \approx 0$), exceed 0.5° even at the middle latitudes (see Table below):

---

* 'tg' ≡ 'tan', 'ctg' ≡ 'cot' — Publisher.

φ	45°	60°	70°	80°	85°
Δ(ΔK)	0°5	0°9	1°4	2°8	5°7

We will determine the conditions under which the deviation of a ship from its course, caused by the error in determining the ship's location on the compass correction, will be insignificant.

Let Point P in Figure 1 represent the North Pole. The navigator assumes that the ship is located at Point C and, in order to reach destination Π, he should maintain the compass course $KK_c = < PC\Pi - \Delta K_c$, obtained for the actual course (from the angle $PC\Pi$) corresponding to the reckoned location and to the compass correction ΔK_C computed for the coordinates of the same point C. Actually, he should have made use of the compass course $KK_M = < PM\Pi - \Delta K_M$ which can not be computed, because the coordinates of the actual location (point M), necessary for determination of the proper course (angle $PM\Pi$) and the proper compass correction (ΔK_M) are not known.

The error in a ship's course caused by the discrepancy between the computed and the actual location is equal to the difference between these compass courses, $KK_c - KK_M$, and the deviation of the ship's course from the point of destination caused by this error will be $\Delta\Pi = (KK_c - KK_M) \sin D = [(< PC\Pi - < PM\Pi) - (\Delta K_c - \Delta K_M)] \sin D$ (3), where $D = M\Pi \approx C\Pi$ represents the angular distance from the location of the ship to the point of its destination, measured along the arc of a great circle.

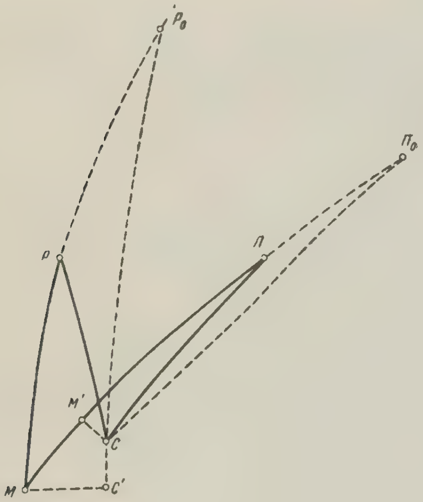


Fig. 1.

In order to compute the difference $(< PC\Pi - < PM\Pi)$ we will extend the meridian MP of the actual location to the point P_0, and the great circle $M\Pi$ to point Π_0, so that $CP_0 = C\Pi_0 = 90°$. With the error on the order of MC^2, which is smaller than 1' at $MC < 1°$, we may take $< PM\Pi = < P_0C\Pi_0$, and we will obtain the desired difference directly from Figure 1, as:

$$(<PC\Pi - <PM\Pi) = <PC\Pi - <P_0C\Pi_0 = <PCP_0 - <\Pi C\Pi_0 = \Delta w \operatorname{tg} \varphi - \Delta m \operatorname{ctg} D \tag{4}$$

where $\Delta m = CM'$ shows the length of a perpendicular line from C to $M\Pi$, which length can be found from the quadrangle CC'MM'C, in which $MC' = \Delta w$, $CC' = \Delta\varphi$, and $C'M \perp PM$ and $CC' \perp C'M - \Delta m = \Delta w \cos(IK) - \Delta\varphi \sin(IK)$.

After substituting in Expression (3) the error $\Delta(\Delta K)$ from (2) for the difference $(\Delta K_c - \Delta K_M)$, and after further substituting of (4) for the difference $(< PC\Pi - < PM\Pi)$, the factor $\Delta w \tan\varphi$ will be cancelled, and the deviation of the ship's course from the point of destination will be expressed by:

$$\Delta\Pi = \Delta n \operatorname{tg} h \sin D - \Delta m \cos D \tag{5}$$

The first product indicates that the heavenly body observed in order to determine the course correction may be chosen on any azimuth, but its elevation should be low. The second product is equal to zero if the determined location lies on the great circle $M\Pi$ connecting the actual location with the destination point, and particularly if the determined location was chosen on the line parallel to the ship's course. In such a case the navigator may compute the compass correction by making use of the coordinates of the point determining this line, despite the fact that the ship's location may not be determined until a second line of direction, intersecting the first one, is found.

The polar region, where the influence of coordinate errors on compass correction is particularly great (see (2)), is no exception to the rule. It is important to measure the azimuth of the celestial body and the ship's course from the same meridian, close to the actual one (MC < 1°).

Besides helping to find the course, the compass correction is needed for determining location by radio signals ([2] p. 303). In the event of a substantial error in determining a location by dead reckoning, the radio signals will carry a systematic error Δ(ΔK), exceeding (see the Table above) 0.3°, which is the allowable accidental error (see [2] p. 315). It can be shown that the error Δ(ΔK) in the compass correction caused by the erroneous determination of a location does not introduce an error into the position determined by the radio signals, if between making the sightings on the heavenly body and receiving the signals the course of the ship is computed in the usual manner, and if the heavenly body is observed at a slight elevation.

It should be noted that an increase in the elevation of the heavenly body would be desirable if it were possible to reach the vertical position of the plane of the sextant when observing the heavenly body (as is done with a universal instrument at the azimuth determinations onshore). In such a case, the observation of no less than three bodies would allow us to make an independent determination of the location and the direction, with an equal accuracy at all the latitudes (for more details regarding determination of a position and direction by measuring the difference in stellar azimuths, see [4]).

The above leads us to the following conclusions:

1) The influence of an error in longitude for the determination of a ship's course, which error becomes larger with the increase in the latitute, may be eliminated at all the latitudes.

2) We cannot agree with most texts on navigational astronomy, which recommend that the selection of the heavenly bodies used in determining compass correction be limited to those near the pole (e. g., see [1], p. 282) or near the first vertical (e. g., see [3], p. 488). Slight elevation of a body observed at any azimuth assures a negligible contribution to the error in the ship's position, and diminishes the importance of instrumental errors.

3) The direction line, when parallel to the course of the ship, not only makes it possible to correct the determination of the ship's position (see [1], p. 225 or [3], p. 436), but also enables us to determine the compass correction for a given course even when the exact position of the ship is not known.

Received October 17, 1956

LITERATURE CITED

[1] A. P. Belobrov, Navigational Astronomy [In Russian] (Hydrometeorology Press, Moscow, 1954).

[2] K. S. Ukhov, Navigation [In Russian] (Maritime and River Press, Moscow, 1954).

[3] B. P. Khliustin, Navigational Astronomy, "Maritime Transport," (Moscow, 1948).

[4] A. B. Marinbakh, Astron. J. (USSR) 28, 274 (1951).

THE DETERMINATION OF AZIMUTHS AT LAPLACE POINTS

B. I. Vlasov

Conditions essential for determining azimuths at Laplace points with equal accuracy at all latitudes are ascertained and the expression (1) is derived for the error of the azimuth at Laplace points due to inaccuracy in the astronomical coordinates. It is found that A. B. Marinbakh's "new" method [2, 3] does not differ from the method proposed earlier by Niethammer, based on the simultaneous determination of the astronomical quantities of Laplace's equation [1]. Application of Formula (1) to the case of separate determination of azimuth and coordinates leads to the conclusion that the advantage of Black's method [4] lies not in the calculation of the azimuth from geodetic coordinates but in eliminating observations of Polaris.

As is well known, it has long been thought that the azimuths at Laplace points are unavoidably burdened with errors of the astronomical coordinates, as a result of which their accuracy decreases with increasing latitude. In the forties it was shown to be possible to obtain Laplace azimuths with the same accuracy at all latitudes (see, for example, [1]).

However, even after the discovery of such a possibility there appeared in print works [2, 3, 4] also devoted to the task of obtaining Laplace azimuths with an accuracy not dependent on the latitude. The very content of these papers, as well as the fact that they have not met with objections up to the present time, testifies to the absence of clarity on this question.

In spite of the differences in the works cited, their authors have a basic idea in common: since, on the one hand, inaccuracies in the astronomical coordinates appear to be the cause of increasing error in azimuth at the Laplace points proportional to tan φ, and, on the other hand, in the process of obtaining the azimuth at a Laplace point the calculation of the astronomical azimuth (i. e., the reduction of the observational results to the astronomical zenith) is an intermediate step, either this stage should be omitted, reducing the observations directly to the geodetic zenith, or an "intermediate point" should be selected having coordinates that are not burdened with error of the astronomical determinations.

Black [4] chose the first course, and A. B. Marinbakh [2, 3] the second, adopting as an intermediate point the so-called "calculated" zenith, defined by the approximate ("calculated") values of the coordinates λ_0, φ_0 considering in this, that "the advantages of the calculated azimuth* over the astronomical, geodetic and Laplace azimuths are based on the fact that the position of the calculated zenith is precisely fixed on the celestial sphere" (without error, so to speak), and "the dependence of the accuracy of the Laplace azimuth on latitude of the place of observation and on errors of the astronomical longitude determination is governed only by the adopted method of obtaining it from the astronomical azimuth" [3], no matter what method the latter is determined by.

---

* I. e., referred to the calculated zenith. — B. V.

But two numbers, whether they are chosen arbitrarily or whether they are obtained as a result of measurements define the position of a point on the surface with complete "accuracy" in the sense that they determine not a quantity, not a region, but a single point. For this reason the position of the astronomical zenith given by coordinates obtained from observations is as "accurately fixed on the celestial sphere" as the calculated one; the error of one or the other consists in the fact that it does not coincide with the true zenith of the place of observation. Consequently, the calculated azimuth does not have any advantage in principle over the astronomical.

For this reason — the precisely fixed relative positions — the reduction of the azimuth from the astronomical zenith to the geodetic must be effected without error and, therefore, by eliminating this stage in obtaining the Laplace azimuth it is also impossible to attain any advantages in principle.

Thus, it is not the elimination of astronomical azimuths as an intermediate step that is essential for obtaining equally accurate Laplace azimuths at all latitudes, but certain other factors which we shall try to ascertain below.

Since the decreasing accuracy of the Laplace azimuths with increasing latitude is a result of errors in the astronomical coordinates, it is possible to neglect all other errors of determining the astronomical azimuth. However, in the case of simultaneous determinations of the azimuth and coordinates, such neglect implies that the coordinates obtained simultaneously with the azimuth are considered to be without error, which would violate the generality of the subsequent statements.

For this reason we assume that all the measurements in the process of determining the azimuth are derived, not from the true astronomical zenith Z^t, but from some other point Z' defined by those values of the coordinates λ', φ' which would be obtained by determining them simultaneously with the azimuth.

Taking into account in this way the final result of the whole combination of observational errors, we shall, however, neglect them in each individual case, considering that each element is in error only in the sense that it is measured from Z' and not from Z^t.

We shall call the point Z' the "azimuthal" zenith.

The zenith defined by the values of the coordinates λ, φ obtained from special observations, with which the calculation of the azimuth is carried out, let us call the "coordinate zenith". Of course, both zeniths, "azimuthal" and "coordinate" are astronomical, i. e., defined by the direction of gravity, but, because of unavoidable inaccuracy in determining this direction, from different sets of observations we obtain different values of the coordinates, more or less differing from the true ones.

If the position of the zenith is determined from the same observational data as the azimuth, then of course it is not necessary to differentiate the "coordinate" zenith from the "azimuthal".

Let the horizontal angle Q' between the star σ and the terrestrial object M (Figure 1) be measured at Z' at the instant of time T (in hours). Knowing the coordinates of the star (α, δ), the correction in hours relative to the time of the starting meridian U and the coordinates of the zenith (λ, φ), determined from special observations, let us compute the azimuth of the star σ as an angle of the triangle $PZ\sigma$:*

$$\operatorname{tg} a_\sigma = \frac{\operatorname{ctg}\delta \sec\varphi \sin(T+u+\lambda-\alpha)}{1-\operatorname{ctg}\delta \operatorname{tg}\varphi \cos(T+u+\lambda-\alpha)} .$$

To determine the azimuth of M it is further necessary to add to a_σ the measured horizontal angle, but let us take account of the fact that the angle Q' in actuality is measured from the "azimuthal" zenith Z', while a_σ refers to the "coordinate" zenith Z.

The error of the astronomical azimuth of the star σ due to inaccurate knowledge of the coordinates, i. e., arising from the fact that it was calculated not with the true values of the coordinates Λ, Φ but with the λ, φ obtained from the observations, is equal to

$$\delta a_\sigma = a_\sigma - a_\sigma^t = \delta\lambda \sin\varphi - (\delta\lambda \cos\varphi \cos a_\sigma - \delta\varphi \sin a_\sigma) \operatorname{ctg} z_\sigma$$

(see [5] p. 180), where:

---

* 'tg' ≡ 'tan', 'ctg' ≡ 'cot' — Publisher.

$$\delta\lambda = \lambda - \Lambda; \quad \delta\varphi = \varphi - \Phi$$

are the true errors of the astronomical coordinates, obtained independently of the azimuth.

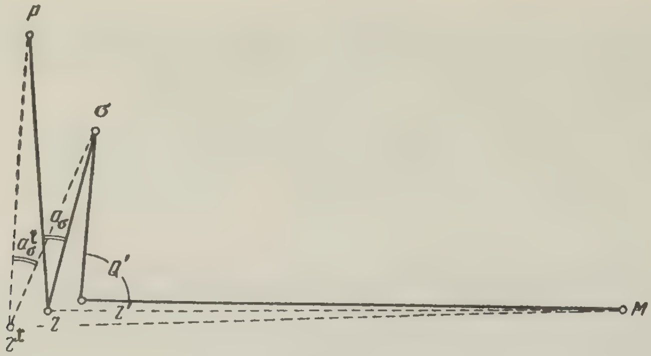


Fig. 1.

The error of the horizontal angle, issuing from the fact that it is measured not from Z^t but from Z, is

$$\delta Q = Q' - Q^{t} = (a'_M - a'_\sigma) - (a^t_M - a^t_\sigma) = (a'_M - a^t_M) - (a'_\sigma - a^t_\sigma) =$$
$$= (\delta\lambda' \cos\varphi \cos a_\sigma - \delta\varphi' \sin a_\sigma) \operatorname{ctg} z_\sigma - (\delta\lambda' \cos\varphi \cos a_M - \delta\varphi' \sin a_M) \operatorname{ctg} z_M,$$

where $\delta\lambda' = \lambda' - \Lambda$, $\delta\varphi' = \varphi' - \Phi$ are errors in the coordinates of the "azimuthal" zenith.

The error in the astronomical azimuth of the object M arising from inaccurate knowledge of the coordinates, or more precisely, from the noncoincidence of the "coordinate" and "azimuthal" zeniths with the true one, equals

$$\delta a_M = \delta a_\sigma + \delta Q.$$

The following expression represents the true value of the Laplace azimuth:

$$A^t_M = a^t_M - (\Lambda - L) \sin\Phi + \frac{(\Lambda - L)\cos\Phi \cos a^t_M - (\Phi - B) \sin a^t_M}{\operatorname{tg} z_M}.$$

Expressing a^t_M, Λ, Φ through measured values and corrections, we obtain

$$A^t_M = a_M - \delta a_M - (\lambda - \delta\lambda - L) \sin\varphi +$$
$$+ \frac{(\lambda - \delta\lambda - L)\cos\varphi \cos a_M - (\varphi - \delta\varphi - B) \sin a_M}{\operatorname{tg} z_M}.$$

Substituting here the value of δa_M and making simple transpositions we obtain:

$$A^t_M = a_M - (\lambda - L)\sin\varphi + \frac{(\lambda - L)\cos\varphi\cos a_M - (\varphi -) \sin aB_M}{\operatorname{tg} z_M} -$$
$$- \left\{ \frac{(\delta\lambda' - \delta\lambda)\cos\varphi\cos a_\sigma - (\delta\varphi' - \delta\varphi)\sin a_\sigma}{\operatorname{tg} z_\sigma} - \frac{(\delta\lambda' - \delta\lambda)\cos\varphi\cos a_M - (\delta\varphi' - \delta\varphi)\sin a_M}{\operatorname{tg} z_M} \right\}.$$

The term in curly brackets represents the error of the Laplace azimuth arising from inaccurate knowledge of the astronomical coordinates, or, more correctly, from the noncoincidence of the "azimuthal" and the "coordinate" zeniths with the true one.

Taking into account the values $\delta\lambda$, $\delta\varphi$; $\delta\lambda'$ and $\delta\varphi'$, let us rewrite the expression for this error:

$$\delta A_M = \frac{(\lambda' - \lambda)\cos\varphi\cos a_\sigma - (\varphi' - \varphi)\sin a_\sigma}{\operatorname{tg} z_\sigma} - \frac{(\lambda' - \lambda)\cos\varphi\cos a_M - (\varphi' - \varphi)\sin a_M}{\operatorname{tg} z_M}. \tag{1}$$

From Formula (1) it is obvious that the error of the Laplace azimuth resulting from inaccurate knowledge of the coordinates depends not on the errors of the astronomical coordinates λ, φ, but on the differences between the coordinates of the "azimuthal" and "coordinate" zeniths $(\lambda' - \lambda)$ and $(\varphi' - \varphi)$.

This error, as is easily seen, represents the difference between the angle Q', measured at the zenith Z', and its value Q, which corresponds to the zenith Z, and is due to the fact that the results measured at Z' refer to Z*.

From this it follows that the astronomical coordinates must be determined simultaneously with the azimuth, since, being distorted by identically the same errors in the observed quantities, the azimuth and coordinates although erroneous in general will nevertheless agree among themselves in the sense that the angle Q' will correspond exactly to the zenith defined by the values of the coordinates obtained, i. e., the concepts of the "coordinate" and "azimuthal" zeniths in this case coincide and the error δA_M will not be present.

When only one star is observed for the determination of the azimuth at a point, it is impossible to speak of an "azimuthal" zenith, as follows from the very definition of this concept. But for each individual observation of this star the Formula (1) will be correct if by λ', φ' are understood the coordinates of the instrumental zenith at the instant of observation of the star (of course, to different observations of the same star there will correspond, in general, different values of λ', φ').

In this case it is impossible to avoid the error δA_M, but Formula (1) points the way to its reduction: the star must be observed with as large a zenith distance as possible, since the Laplace azimuth will be influenced by the errors in the astronomical coordinates in proportion to cot z, i. e., to the same extent as for the reduction of the observational results directly to the geodetic zenith, as was done by Black [4]. The decrease in accuracy of the Laplace azimuth in proportion to tan φ is a specific peculiarity of observations of Polaris, for which $z \approx 90° - \varphi$ and $\cot z = \tan\varphi$.

Now in the method suggested by Black, the essential factor, in spite of the opinion of the author, is not the reduction of the results directly to the geodetic zenith, but the elimination of the observations of Polaris.

In order to avoid unnecessarily complicated formulas and diagrams we will suppose further that the astronomical azimuth is determined from observations of stars on a vertical circle through the object, i. e. by the direct method sufficiently explained in the domestic literature [6, 7, 8].

For this method $a_M = a_\sigma = a$ and Formula (1) takes the form:

$$\delta A = \{(\lambda' - \lambda)\cos\varphi\cos a - (\varphi' - \varphi)\sin a\}(\operatorname{ctg} z_\sigma - \operatorname{ctg} z_M), \tag{2}$$

or

$$\delta A = \{(\lambda' - \lambda)\cos\varphi\cos a - (\varphi' - \varphi)\sin a\}\frac{\sin(z_M - z_\sigma)}{\sin z_M \sin z_\sigma}. \tag{2'}$$

As is evident from Figure 2, δA is measured by the angle $MZ\sigma$ and results from the fact that the object and the star, being at the instant of observation in one "azimuthal" vertical circle $Z'\sigma M$, are found from the "coordinate" zenith Z in the different verticals $Z\sigma$ and ZM.

Let us consider the case where one of the coordinates is determined simultaneously with the azimuth, but the other is found from an independent set of observations.

Let the position of the vertical Z'M be determined on the celestial sphere by observations of the transit of not less than two known stars (see [1] Chapter I).

A) The latitude φ is found from separate observations.

---

*Let us recall that we are ignoring observational errors as pointed out previously.

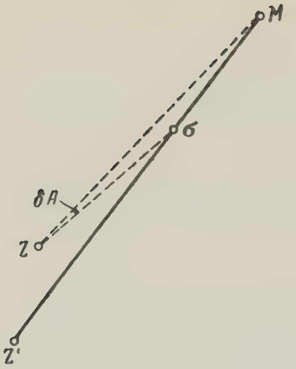


Fig. 2.

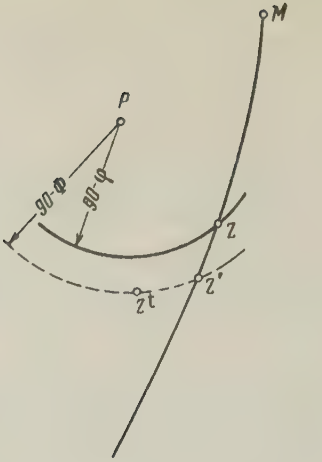


Fig. 3.

If we knew the true value of the latitude Φ and the observations in determining the azimuth and longitude were without error, then, at the intersection of the parallel φ (Fig. 3) with the correct position found for the vertical circle through the object, we would obtain the true zenith Z^t. The intersection of the parallel Φ with the position of the vertical circle Z'M, obtained from the actual observations, gives the zenith Z'. As a matter of fact, we shall know the value of the latitude φ, in general incorrect by the amount $\delta\varphi = \varphi - \Phi$, and the zenith will be defined as the point of intersection Z of the parallel φ with the position of the vertical Z'M actually found from observations. Thus the error in the position of Z depends on the errors of the set of observations from which the azimuth and longitude are derived, as well as on the error of the independently determined latitude. Hence the displacement of the zenith, owing to the latitude error, from the position Z', which is distorted only by errors in the observations used to derive the azimuth and longitude, to the position Z, originates in the vertical circle Z'M (Fig. 3).

B) The longitude λ is found from separate observations.

The position of the zenith is found as the point of intersection of the vertical circle Z'M fixed by observation and the meridian of longitude λ (Fig. 4).

It is easy to see that the displacement of the zenith from the position Z' to Z, due to the error in the independently observed longitude $\delta\lambda = \lambda - \Lambda$, will originate, as in case A, in the vertical circle Z'M.

Now in both cases the displacement of the zenith due to the inaccuracy of the separately observed coordinates originates in the vertical circle Z'M; therefore the error δA will not be present, as is apparent from its geometrical interpretation (Fig. 2).

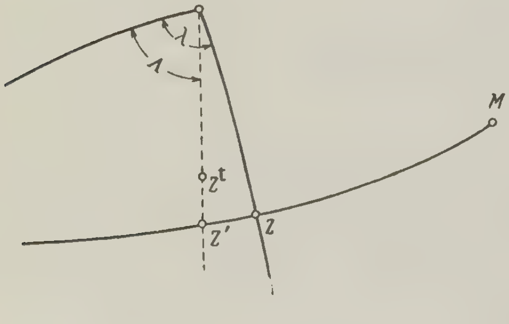


Fig. 4.

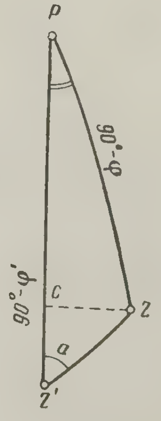


Fig. 5.

The latter fact can also be confirmed analytically.

In fact, let the displacement Z'Z (Fig. 5) lie in the vertical circle through the object. Then from the elementary right triangles PZc and ZZ'c we find

$$cZ = (\lambda - \lambda') \cos\varphi, \qquad cZ = cZ' \operatorname{tg} a = (\varphi - \varphi') \operatorname{tg} a,$$

from which

$$\varphi - \varphi' = (\lambda - \lambda') \cos\varphi \operatorname{ctg} a.$$

Substituting this value of $(\varphi - \varphi')$ in Formula (2) we obtain:

$$\delta A = 0.$$

Now it is possible to extend the conclusion drawn earlier: the Laplace azimuth is wholly undistorted by errors of the astronomical coordinates and as a result it can be obtained with equal accuracy at all latitudes if one or both of the coordinates are determined simultaneously with the astronomical azimuth.

Thus "it is appropriate to raise the question" not " of separating the task of astronomical determination of azimuth at geodetic points from the task of determining the direction of the gravitational vertical" [3], but on the contrary, that of determining the astronomical azimuth and one or both of the coordinates.

Actually, the equations of the form:

$$\Delta a^{\circ} \sin z_i + \zeta^{\circ} \cos z_i - l_i = v_i \qquad \text{([3] Formula (4.3))},$$

from which A. B. Marinbakh determines the correction to the calculated azimuth, are found to be equations for the simultaneous determination of the azimuth and the coordinates, entirely identical with the expressions found earlier by many authors who studied the problem of simultaneous determinations, as for example:

$$x \sin z_j + y \cos z_i = l_i \sin z_i + \lambda_i \qquad \text{([1] Formula (69))},$$

where $x = \Delta a^0 = \Delta a - \Delta u \sin\varphi$, $y = \zeta^0 = \Delta u \cos\varphi \cos a - \Delta\varphi \sin a$, $\lambda_i = v_i$ is a correction to the free term, in both cases equal to the product of sin z_i by the difference between the azimuth of the star a_i, computed with the approximate ("calculated") coordinates λ_0, φ_0, and the approximate value adopted for the azimuth of the terrestrial object a_0.

Particularly surprising is the fact that A. B. Marinbakh, quoting in his work [3] the well known equation for simultaneous determination of Δa, Δu, $\Delta\varphi$ (1.1), speaks of it as different from (4.3).

Perhaps he considers it an essential advantage, distinguishing (4.3) from (1.1), that in the first expression the content of the symbol Δa^0 is not displayed.

Of course, for the calculation of the azimuth at a Laplace point it is not necessary to separate the differential corrections Δa, Δu and $\Delta\varphi$ since they are connected with the unknowns $x = \Delta a^0$ and $y = \zeta^0$ obtained from the equation in the same form as the corresponding final quantities of the Laplace equation (an interesting circumstance, remarked apparently for the first time by Niethammer), just as it is unnecessary to combine the term $(x + y \cot z_M) = (\Delta a^0 + \zeta^0 \cot z_M)$ with a_0, introducing for the sum $a_0 + x + y \cot z_M$ the specific designation "calculated azimuth."

However, if one separates the corrections and calculates the Laplace azimuth from the astronomical azimuth $a = a_0 + \Delta a$ and the coordinates $\lambda = \lambda_0 + \Delta\lambda$ and $\varphi = \varphi_0 + \Delta\varphi$, the errors of the coordinates will not show any effect on the result, i. e., the accuracy of the geodetic azimuth in this case too will not depend on the latitude, owing to the interdependence and concordance of all the errors. This is easily shown by simple calculations.

Now in the method which was suggested in our literature by A. B. Marinbakh the essential factor is precisely the simultaneity of obtaining the corrections to the astronomical azimuth and coordinates and by no means any magical properties of the "calculated" zenith, as the author of the suggestion seems to think.

In conclusion let us note the inaccuracy of the categorical statement that "all systematic errors in determining the astronomical longitude which grow with increasing latitude of the place of observation" [3] enter into the azimuth at a Laplace point in the case where the astronomical azimuth is determined by the direct method and the coordinates are found by independent observations.

It is also shown that there exists a possibility, implied by the formula (2), of obtaining the geodetic azimuth with equal accuracy at all latitudes. Actually, in observing two stars on different sides of the zenith at equal distances from it, we completely free ourselves from the effect of errors in the coordinates, because δA for the azimuth ($a + 180°$) changes sign. In observing many stars on one vertical circle, the rough condition of equality of zenith distances in pairs is replaced by the condition:

$$\sum_{a} \frac{\sin(z_M - z_\sigma)}{\sin z_\sigma} = \sum_{a+180^\circ} \frac{\sin(z_M - z_\sigma)}{\sin z_\sigma}$$

or approximately:

$$\sum_{a} \operatorname{ctg} z_\sigma = \sum_{a+180^\circ} \operatorname{ctg} z_\sigma$$

As we see, in this case the reference to the formula

$$m_{A_M}^2 = m_{a_0}^2 + (m_{a_0}^2 + m_T^2)\operatorname{tg}^2\varphi \quad ([3]\,(2.2))$$

is unfounded.

Received July 20, 1956

LITERATURE CITED

[1] Th. Niethammer, Die genauen Methoden der Astronomischgeographischen Ortsbestimmung (Basel, 1947) Chapter V, Sections 5, 7.

[2] A. B. Marinbakh, Doklady Akad. Nauk SSSR 64, 791 (1949).

[3] A. B. Marinbakh, Astron. J. (USSR) 28, 263 (1951).

[4] A. Black, Empire Survey Review, 11, 177 (1951).

[5] N. Tsinger, A Course in Practical Astronomy [in Russian] (Moscow) p. 160.

[6] N. A. Beliaev, Astron. J. (USSR) 30, 196 (1953).

[7] N. N. Goriachev, Trudy Tomsk. Univ. 86, 88 (1934).

[8] N. N. Goriachev, Astronomical Determinations of the Kuzbass and Gornaia Shoriia Stations [in Russian] (Leningrad, 1934).

THE INFLUENCE OF DIURNAL ABERRATION ON THE ACCURACY OF DETERMINATION OF LATITUDE FROM OBSERVATIONS OF POLARIS

V. F. D'iakonov

An analysis is given of the influence of diurnal aberration on the equatorial coordinates (α and δ) of Polaris, and of its effect on the accuracy of latitude determination from observations of Polaris in connection with studies of Class II astronomical positions.

On the basis of analytical investigation and numerical tables, the author has shown that the influence of diurnal aberration on declination and right ascension can give rise to a total error (for $t < 90°$) in latitude on the order of $\pm\ 0.''5$. For this reason, in reducing observational results it is necessary to take into account the effect of the diurnal aberration on the geographical latitude determined through observations of Polaris.

In 1954 the scientist A. V. Burkevich, of the Novosibirsk Geodetic Institute, introduced a suggestion regarding the neglect of the effect of diurnal aberration in determining geographical latitude from observations of Polaris. In existing courses of practical and spherical astronomy, it is not always possible to find sufficiently clear statements indicating in which cases the effect of diurnal aberration on the final results should be taken into account, and when such an effect can be discounted.

More complete instructions on this question are given in Professor M. K. Venttsel's course in spherical astronomy [1]. Analyzing the formulas

$$\begin{aligned} \Delta\alpha &= 0''.32 \cos\varphi \sec\delta \cos t, \\ \Delta\delta &= 0''.32 \cos\varphi \sin\delta \sin t, \end{aligned} \tag{1}$$

which take into account the effect of diurnal aberration on the equatorial coordinates of stars, Professor M. K. Venttsel' writes: "It is particularly notable that because of the factors sec δ and sin t, the influence of diurnal aberration, in general negligible, attains a more significant magnitude for stars with greater declinations.

For this reason the effect of diurnal aberration is generally taken into account only for stars with declinations of 80° to 90°, in particular for Polaris (αUMin)".

In accordance with existing Instructions [2] and Directions for cameral reduction of astronomical observations [3], the method of determining the latitude from the zenith distance of Polaris is used in astronomical work at Class II points.

According to instructions for work of Class II, the mean square error of latitude, derived from a comparison of separate results, should not exceed $\pm\ 0.''4$. Therefore, in order to insure the accuracy mentioned in the Instruction for cameral reduction of astronomical determinations, it is advisable in reducing observational results to select stellar coordinates from the Astronomical Ephemeris to an accuracy of $0''.1$, and for observations of Polaris to correct α and δ for diurnal aberrations by the formulas:

$$\Delta\alpha = 0''.021 \cos\varphi_0 \cos t_{av} \sec\delta,$$

$$\Delta\delta = 0''.32 \cos\varphi_0 \sin t_{av} \sin\delta,$$

here $t_{av} = T_{av} - (\alpha - u)''$ ([3] p. 49) and φ_0 is the approximate latitude which must be known to an accuracy of 0'.1.

First of all, we shall explain the direct influence of diurnal aberration on the equatorial coordinates of Polaris.

For the right ascension the influence referred to is shown by the formula:

$$\Delta\alpha = 0''.32 \cos\varphi \cos t \sec\delta$$

and for $\delta = 89^\circ$ ($\sec\delta = 57.299$) it is specified in Table 1.

The influence of diurnal aberration on the declination of a star is specified by the analogous Table 2, computed according to the formula

$$\Delta\delta = 0''.32 \cos\varphi \sin t \sin\delta \quad (\sin\delta = 1).$$

TABLE 1

$\Delta\alpha$

t / φ	0°	30°	60°	80°	90°
0°	18″3	15″8	9″2	3″2	0″0
30	15.8	13.7	7.8	2.7	0.0
60	9.2	7.8	4.6	1.6	0.0
80	3.2	2.7	1.6	0·6	0.0

TABLE 2

$\Delta\delta = \Delta\varphi_\delta$

t / φ	0°	30°	60°	80°	90°
0°	0″0	0″2	0″3	0″3	0″3
30	0.0	0.1	0.2	0.3	0.3
60	0.0	0.1	0.1	0.2	0.2
80	0.0	0.0	0.0	0.1	0.1

Regardless of the method of computing the latitude determined by the zenith distance of Polaris, the error of declination $\Delta\delta$ will show a direct influence on the computed latitude [$\varphi = z + \delta$ or $\varphi = 180^\circ - (z + \delta)$].

Besides this, the error of right ascension $\Delta\alpha$ will show a substantial influence on the final derivation of the latitude. To make clear the influence of diurnal aberration in right ascension of the determined latitude we refer to the appropriate reduction formula [4]:

$$r'' = \frac{2\cos\varphi\cos\delta}{\sin\frac{1}{2}(z + z_m)} \; \frac{\sin^2\frac{t}{2}}{\operatorname{arc} 1''}. \tag{2}$$

Differentiating the Expression (2) with respect to the variables r and t we derive

$$\Delta r''_\alpha = \frac{\cos\varphi\cos\delta\sin t}{\sin\frac{1}{2}(z + z_m)\operatorname{arc} 1''}\,\Delta\alpha. \tag{3}$$

Substituting the value $\Delta\alpha = 0''.32 \cos\varphi \sec\delta \cos t \operatorname{arc} 1''$, we find

$$\Delta r''_\alpha = \frac{0''.32\cos^2\varphi\sin t\cos t}{\sin\frac{1}{2}(z + z_m)},$$

or taking into account $\Delta r_\alpha = \Delta\varphi_\alpha$, we will have

$$\Delta\varphi_\alpha = \frac{0''.16\cos^2\varphi\sin 2t}{\sin\frac{1}{2}(z+z_m)}. \qquad (4)$$

The computed values of the errors $\Delta\varphi_\alpha$ are specified by the following Table 3.

TABLE 3

φ	t / $\frac{1}{2}(z+z_m)$	0°	30°	60°	80°	90°
0°	90°	0″.0	0″.2	0″.2	0″.2	0″.0
30	60	0.0	0.2	0.2	0.2	0.0
60	30	0.0	0.1	0.1	0.1	0.0
80	10	0.0	0.0	0.0	0.0	0.0

The simultaneous influence of diurnal aberration in right ascension and in declination on latitude determination from observations of Polaris is specified by the following Table 4.

TABLE 4

$\frac{1}{2}(z+z_m)$	t / φ	0°	30°	60°	80°	90°
90°	0°	0″.0	0″.4	0″.5	0″.5	0″.3
60	30	0.0	0.3	0.4	0.5	0.3
30	60	0.0	0.2	0.2	0.3	0.2
10	80	0.0	0.0	0.0	0.1	0.0

From the table shown, it is seen that in latitude determination from observations of Polaris with an accuracy $\epsilon_\varphi = +0''.4$ (for work of Class II) one must not neglect the influence of diurnal aberration.

In clases where latitude is determined by the approximate method with the help of corrections I, II and III shown in the special tables of the Astronomical Ephemeris, it is not essential to take into account the influence of diurnal aberration.

Received July 20, 1955

LITERATURE CITED

[1] M. K. Venttsel', Spherical Astronomy [in Russian] (Geodezizdat, 1952), p. 210.

[2] Instructions for Astronomical Determinations of Points of Classes I, II, III and IV [in Russian] (Geodezizdat, Moscow, 1948), p. 15.

[3] Directions for Cameral Reduction of Astronomical Determinations [in Russian] (Redbiuro GUGK, 1939), pp. 48-50, 67-79.

[4] K. A. Tsvetkov, Practical Astronomy (1951), p. 146.

PHOTOGRAPHING STARS BY MEANS OF AN IMAGE CONVERTER TUBE

P. V. Shcheglov

We present a photograph of the galactic cluster M39 taken with infrared rays by means of an image converter tube.

In the summer of 1955 an attempt was made to photograph stars by means of an image converter attached to the camera ($d = 64$ cm, $d/f = 1/1.4$) of the Crimean Astrophysical Observatory of the Academy of Sciences of the USSR. An image-converter tube with an oxygen-cesium cathode was employed. The cathode of the converter was cooled with dry ice. In front of the cathode we mounted a filter which did not transmit wavelengths below 8000 A. The exposure did not last for more than a few minutes because of the intense background of hydroxylic radiation in the night sky and the large aperture of the telescope employed.

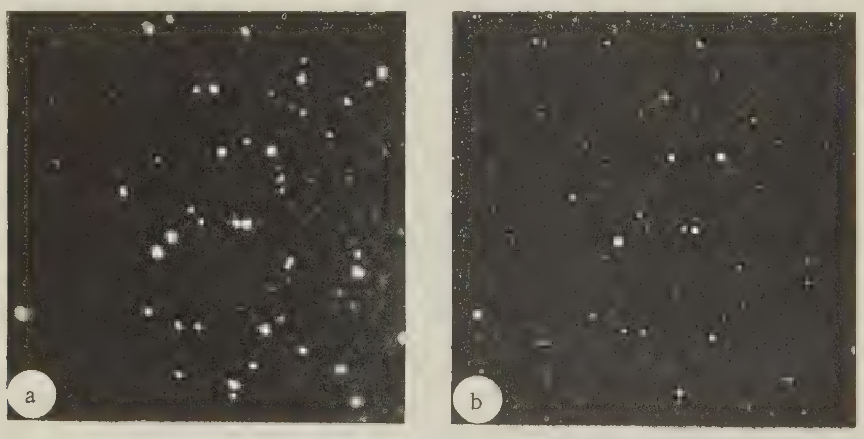


Fig. 1. a) Photograph of cluster M39 by infrared rays; b) photograph of M39 by photographic rays.

Figure 1,a shows a photograph of the galactic cluster M39 (NGC 7092, $\alpha\,(1950.0) = 21^{h}30^{m}4$, $\delta\,(1950.0) =$ $= \pm\,48°13'$) by infrared rays with exposure 1 min.

Figure 1,b shows a photograph of the same cluster, taken with the Schmidt telescope of the Biurakan Astrophysical Observatory ($d = 48$ cm, $d/f = 1/3.8$), exposure 30 min. The photographs are on the same scale and have the same orientation. Comparison of the photographs shows that the use of infra-red photography results in clearly defined stars with large color indices.

The use of lower-aperture telescopes will permit a longer exposure and the observation of weaker stars.

The photographs given here were on show at the 1955 Dublin meeting of the International Astronomical Union, but their publication in the Proceedings of the Meeting was impossible for technical reasons.

I express my profound gratitude to those in charge of the Crimean Astrophysical Observatory of the Academy of Sciences of the USSR for the facilities provided.

P. K. Shternberg State Astronomical Institute

Received December 11, 1956

HIGH VOLTAGE SOURCES FOR SUPPLYING IMAGE CONVERTER TUBES

P. V. Shcheglov and V. G. Kurt

A description is given of high voltage mains and battery power supplies for image converter tubes. The sources described were used during observations in 1955-1956.

For supplying image converters we require a voltage of several tens of kilovolts and a very small current. To obtain such voltages by direct transformation of current at 50 cps is possible, but it involves difficulties in the construction of the step-up transformer. Other snags in using such a high voltage source are connected with the need to employ condensers of relatively high capacitance in the rectifier filter. These are required because of the low frequency of the rectified voltage, and also because of the low back resistance of the selenium stacks usually employed in rectifiers of this type. Hence the filter of such a rectifier stores up a considerable amount of electricity which may be dangerous for the observer if he chances to come in contact with the current-carrying parts.

High voltage, however, can also be obtained by other methods. One of such methods is the conversion of voltage by means of an inductance coil. As we know, energy stored by the magnetic field of a coil when current is passed through it is equal to $1/2\ Li^2$, where L is the inductance of coil, and i is the current flowing through it. If the current flowing through the coil is suddenly switched off, the energy stored in the magnetic field is changed into electrical energy which charges the spurious inter-turn capacitance of the coil to a certain voltage U. The energy stored in a condenser charged to voltage U is equal to $1/2\ CU^2$, where C is the capacitance of the condenser. Hence, when the current flowing through the coil is suddenly switched off, there arises on its terminals a voltage pulse, the amplitude of which is determined from the equivalence of the energy stored in the magnetic field and the energy stored in the spurious capacitance of the coil:

$$^1/_2\,Li^2 = {}^1/_2CU^2; \quad U = i\sqrt{\frac{L}{C}}.$$

With a fixed value of current flowing through the coil, the amplitude of the pulse will depend on the ratio of the inductance of the coil to its spurious capacitance. With a coil of suitable construction this ratio may be made sufficiently large, and the pulse amplitude will reach several kilovolts.

The energy stored in the spurious capacitance of the coil produces damped oscillations in the circuit formed by the coil and its spurious capacitance. These damped oscillations can be rectified by means of the usual rectifier circuit to give a constant voltage equal to the maximum amplitude of the oscillations (when the load resistance is large enough).

A set-up of this kind was constructed in the Radioastronomy Department of GAISh* in 1955. The circuit is shown in Figure 1. The assembly consiste of a multivibrator, blocking and unblocking a tube, to the anode of which the inductance coil is connected. As a coil we used the ignition coil of the automobile "Pobeda". The voltage supplied by the power source is regulated by a change of bias on the grid of a 6P7S tube, which is achieved by altering the value of its cathode resistance. The rectifier circuit is a voltage-doubling circuit. Because of the high operating frequency (70 kc) the capacitance of the condensers in the rectifier circuit is quite small; therefore the operation of such a supply source is quite safe. By supplying the filaments of the rectifier tubes from a separate well-insulated transformer and not from additional turns on the inductance coil, we avoid the introduction of additional damping and reduction of the rectified voltage. The whole rectifier circuit and the transformer for the filament supply are embedded in a mixture of rosin and paraffin wax, which protects it from dampness. The size of the supply unit is 250 x 200 x 200 mm, its weight is 5 kg. An oscillogram of the voltage on the anode of the 6P7S tube is given in Figure 2. The current given by such a power source can reach 10^{-5} amp, which is sufficient for supplying circuits with high-ohmic dividers. The output voltage is controlled by means of a S-96 voltmeter or a microammeter with an additional resistance of 20 kM Ω. The source may be supplied directly from the 110/220 V ac mains; if the mains voltage fluctuates a lot, it is advisable to employ a ferro-resonance stabilizer, if an output voltage varying by less than a few percent is required.

---

* Shternberg State Astronomical Institute.

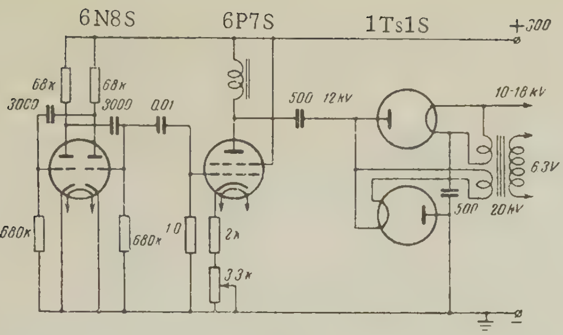


Fig. 1. Circuit of main source.

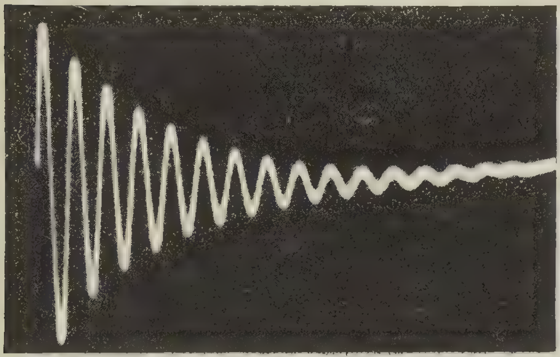


Fig. 2. Form of voltage on anode of 6P7S tube.

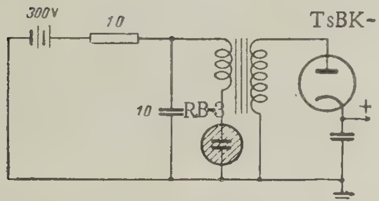


Fig. 3. Circuit of battery source.

For work in the field we designed a second assembly for a high-voltage supply. The circuit of the power supply is shown on Figure 3. The battery GB-300 charges a condenser through a resistance. When the voltage on the condenser reaches 300 V, it discharges through a discharge switch RB-3, in the circuit of which is connected the primary winding of the transformer. The high voltage is taken from the secondary winding of the transformer. As kenotrons we used filamentless kenotrons TsBK-1. They do not require filament voltage, which greatly simplifies the design of the rectifier circuit. Such kenotrons can also be employed in the rectifier of the mains source described above. Good insulation of the high-voltage section of the battery power supply was ensured by embedding it in a mixture of rosin and paraffin wax. As condenser for the rectifier filter we used the capacitance of the screened cable taking the voltage to the image converter tube. The power supply consumes 0.3 ma at 300 V. The sources described above were used by the authors for astronomical observations during 1955-1956.

P. K. Shternberg State Astronomical Institute

Received December 11, 1956

A BOLOMETRIC PYRHELIOMETER FOR ABSOLUTE MEASUREMENTS OF DIRECT SOLAR RADIATION

P. V. V'iushkov and E. A. Ridel'

A description is given of the construction of a pyrheliometer based on the bolometric principle, the radiation receiver having the form of a plane single-layer coil. The bolometric pyrheliometer has several advantages over the Angstrom compensation pyrheliometer. Preliminary investigations showed good agreement between the results of measurements. The instrument can be used for absolute determinations of solar radiation.

Direct solar radiation, which is of great interest from the practical as well as the theoretical point of view, is studied by means of several actinometric instruments which invariably require special comparison with primary intruments (pyrheliometers) in order to determine the coefficients for converting their readings to calories.

The primary pyrheliometers in use at the present time are the Angstrom compensation pyrheliometer, which is widely employed in Europe, and the Abbot water-flow pyrheliometer in America. A comparison of the results obtained with the Angstrom and Abbot pyrheliometers, however, revealed a systematic difference of the order of 3-4% between them. Thus at present there does not exist a unique scale of measurement for solar radiation, and the problem of constructing a primary pyrheliometer for its establishment is of real importance.

In 1946 P. V. V'iushkov suggested a new type of pyrheliometer based on the bolometric principle. The bolometric pyrheliometer is a balanced resistance bridge, one arm of which, acting as the radiation receiver, has a high temperature coefficient of resistance. This allows the bridge to be balanced by altering the current flowing through it, without changing the resistance of the other arms.

The temperature of the radiation receiver, and hence its resistance also, is a certain function of the energy received by it. Thus, with other conditions equal, equality of the receiver resistances will signify equality of energies received by it under different conditions of irradiation.

If the radiation receiver is subjected to the action of solar radiation, the energy received by it from the sun and the current, in calories per minute, can be expressed by the formula:

$$E_1 = IS\sigma + 14.35\, ri_0^2,$$

where I is the intensity of solar radiation, S is the area, σ is the absorptive power and r is the resistance of the receiver, i_0 is the current flowing through it.

With the receiver in the dark the energy will be equal to

$$E_2 = 14.35\, ri_d^2,$$

where i_d is the current.

When the bridge is balanced the receiver resistance is the same, and hence the energy E_1 must be equal to E_2. From this we obtain the formula for determining the intensity of solar radiation

$$I = \frac{14.35r}{S\sigma}\left(i_d^2 - i_0^2\right). \qquad (1)$$

Since it is possible to determine independently the constants S, σ and r in Formula (1) the instrument can be used for absolute measurements of direct solar radiation.

The main scheme of the instrument is shown in Figure 1. The instrument is composed of two units. The radiation receiver 1, inserted in the first arm of the bridge, is made from enameled copper wire about 0.05 mm thick in the form of a plane one-layer coil with a receiving surface S of the order 0.5-1.0 cm² and resistance r about 10 Ω. The receiving surface of the coil is coated with a fine layer of soot which gives an absorptive power σ in the region of 0.98.

The coil, fixed by means of fine wires to an ebonite ring, is mounted in the lower end of a metal tube which carries a number of internal diaphragms, forming a cone with angle 5°. The diameter of the outer diaphragm subtends an angle of 10° at the center of the coil. The tube has a sighting device for tracking the sun and is mounted on a parallactic base.

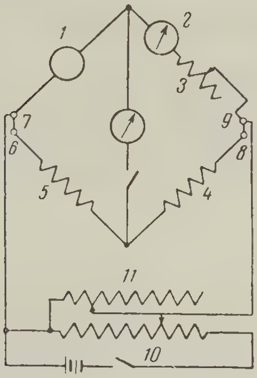


Fig. 1.

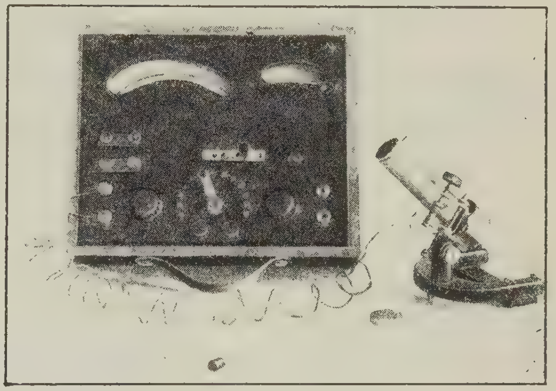


Fig. 2.

The second arm of the bridge consists of a milliammeter 2 and a resistance box 3 which allows the resistance of the arm to be changed by 0.1 Ω.

The other two arms 4 and 5 of the bridge are made from manganin wire and have equal resistances of the order 10 Ω.

Contacts 6, 7, 8, 9 are brought out to the external panel of the second unit to allow an interchange of resistances 4 and 5 for a check of their equality.

When the bridge is balanced, the milliameter in the second arm shows the current flowing through the radiation receiver, while the reading on the resistance box gives the resistance of the first arm.

The rheostats 10 and 11 ensure that the current flowing through the detector can be varied smoothly.

A general view of the apparatus is depicted in Figure 2.

When solar radiation is being measured the bridge is balanced by means of the box 3 and rheostats 10 and 11 with the receiver tube covered. When the tube is opened the receiver is subjected to irradiation, and the bridge is balanced by changing the resistance of rheostats 10 and 11 with a constant reading on the box.

Knowing the value of the current passing through the receiver in the first (i_d) and second (i_0) cases, and also the resistance r of the receiver, we can calculate the value of the radiation from Formula (1).

This bolometric pyrheliometer has a number of advantages over the Angstrom compensation pyrheliometer, the most important being: a great reduction in the edge effect, the circular form of the receiver, the employment of the receiver itself as a temperature meter and the fact that the high sensitivity of the bolometric arrangement permits the use of a coarser galvanometer.

In 1949 preliminary measurements of the solar radiation were made with a model of the bolometric pyrheliometer with two different receivers.

In 1953 Iu. A. Skliarov [4] conducted a series of measurements with a prototype of the bolometric pyrheliometer with three different receivers.

A comparison of these results gives a value for the solar radiation from 35 sets of measurements 1.4 % higher on the average than the results obtained from parallel measurements with an Angstrom pyrheliometer No. 258 with the mean square error of one observation ± 0.008 cal/cm^2· min, i. e., 0.6 %.

From the good agreement in results obtained with one receiver, as well as with different receivers, we may hope to obtain a new, truer scale for the measurement of direct solar radiation.

N. G. Chernyshevskii Saratov State University　　Received December 3, 1956

LITERATURE CITED

[1] N. N. Kalitin, Actinometry [in Russian] (Hydrometeorol. Press, Moscow-Leningrad, 1938).

[2] A. Unsold, Physik der Sternatmospharen [Russian translation] (IL, 1949).

[3] Trudy Saratovskogo Otdeleniia Vsesoiusnogo Astronomo-Geodezicheskogo Obshchestva, 1, 95 (1947).

[4] Iu. A. Skliarov, Scientific Year Book of the Saratov State University [in Russian] (Saratov, 1955), p. 633.

LETTERS TO THE EDITORS

A NOTE ON THE PAPER "LUMINOSITY FUNCTION FOR STARS OF THE MAIN SEQUENCE AND ITS INTERPRETATION" BY A. G. MASEVICH

G. M. Nikol'skii

In a recent paper A. G. Masevich [1] considers the important problem of the evolution of stars of the main sequence. The analysis of the problem is carried out as follows.

Suppose the stars are formed continuously and in a unit time n stars enter the main sequence at the point M_0 (absolute stellar magnitude). Evolution takes place in the same way for all the stars which move down the main sequence. If t is the time of evolution of a star from the point M_0 to the point M then (cf. Equation (7) in [1]):

$$nt = N \int_{M_0}^{M} \varphi(M)\, dM. \tag{1}$$

Here φ is the luminosity function observed at the present time and N is the number of stars in the galaxy. The left hand side of Equation (1) involves the number of stars formed during the time t and the right hand side the number of stars observed at the present time in the interval M_0, M.

Substituting into Equation (1) the expression for $\varphi(M)$, and going over to luminosity L, Masevich obtains (cf. Formula (13) in [1]).

$$22\frac{nt}{N} = \frac{1}{L} - \frac{1}{L_0}. \tag{2}$$

(in fact the powers of L are equal to 0.9 and not 1.0 but this does not affect the argument).

A. G. Masevich then argues as follows: if a star during its evolution loses mass, then using (2) and certain other data it is possible to estimate the number of stars formed during the time t. If the evolution of the star occurs at constant mass then according to [2] its luminosity should increase. However, it follows from (2) that if L_0 is less than L then n is always negative. This absurd conclusion indicates, in the opinion of A. G. Masevich, that evolution can only take place with the loss of mass. Furthermore, A. G. Masevich considers that the negative sign indicates the exit of stars which evolve at constant mass from the main sequence.

General considerations force the reader to doubt the possibility of the solution of the evolutionary alternative in this way.

The right hand side of Equation (1) involves the number of stars observed in the given interval of absolute magnitudes. It is quite clear that this quantity is essentially positive. Formally, A. G. Masevich's mistake can be reduced to the following: in considering the evolution of the second type $M_0 > M$ ($dM < 0$), and the limits of integration were not interchanged on the right hand side of Equation (1).

Thus Equation (2) which is a mathematical consequence of Equation (1) says nothing about the direction in which the evolution of stars proceeds since n is always positive. It follows that all the subsequent conclusions of the author of [1] are a consequence of a mathematical mistake.

LITERATURE CITED

[1] A. G. Masevich, Astron. J. (USSR) 33, 2 (1956).

[2] A. G. Masevich, Soobshch. GAISh, 99, 1 (1955); Astron. J. (USSR) 30, 5 (1953).

A REPLY TO G. M. NIKOL'SKII'S NOTE

A. G. Masevich

G. M. Nikol'skii's note on my paper "Luminosity Function for Stars of the Main Sequence and its Interpretation" [1] is based on a simple misunderstanding.

It is well known that in all possible cases of the evolution of stars at constant mass the luminosity and radius of a star increase. Since the mass remains constant it follows that, in time, the star ceases to obey the mass-luminosity and mass-radius functions characteristic of the main sequence i. e., it "leaves" the sequence. This result was discussed in detail in our previous papers [2, 3]* on the basis of special calculations and does not, as indeed it cannot, be a consequence of Equation (2) (Equations (13) in [1]), as is for some reason considered by G. M. Nikol'skii. The opposite is the case since in [1] the result obtained from a consideration of evolution at constant mass, without the interchange of the limits of integration in the integral (1), is treated (possibly not altogether successfully) as a mathematical expression of the above known fact which is quite trivial and can be omitted without loss since it has no effect on the subsequent conclusions.

The choice between the two alternative evolutionary hypotheses using the observed luminosity function is based in [1] not on the sign but the numerical value of the quantity $\underline{n}$ ($\underline{n}$ is the number of stars being formed, which is necessarily positive).

It was shown in [1] that for evolution at variable mass $n = 5 \cdot 10^8$ stars per galaxy per 10^{10} years in accordance with the estimate of the number of stars formed in O-associations ($10^8 - 10^9$ stars per 10^{10} years) carried out by V. A. Ambartsumian. For evolution at constant mass this number turns out to be higher by one order, i. e., $n \approx 5 \cdot 10^9$ stars.

It is easy to see that in this case this is the minimum value of $\underline{n}$ since we are considering evolution without intermixing and the star is of type A2 whose evolution takes place relatively slowly. If the same calculation is carried out for stars of earlier spectral classes then it turns out that for B-stars $n \approx 10^{10}$ and for O-stars $n \approx 10^{11}$ [4] i. e., comparable with the total number of all the stars in the galaxy which is in fact absurd.

This is the main argument given by us in [1] in support of the inconsistency between the observed luminosity function and the hypothesis of evolution at constant mass (p. 220).

G. M. Nikol'skii, having focussed his attention on the unimportant question of sign, apparently "did not notice" this very important result.

Thus the conclusions of [1] are correct and the concluding paragraph of Nikol'skii's letter is unjustifiable. It only remains to express regret for the misunderstanding which has led to the present note.

---

* It is very surprising how G. M. Nikol'skii can refer to my paper [2] since he could not have seen it as the entire issue of Soobshcheniia GAISh No. 99, although printed, has not as yet been issued by the Moscow University Press because of administrative troubles.

LITERATURE CITED

[1] A. G. Masevich, Astron. J. (USSR) 33, 2 (1956).

[2] A. G. Masevich, Soobshcheniia GAISh 99, 1 (1955).

[3] A. G. Masevich, Astron. J. (USSR) 30, 5 (1953).

[4] A. G. Masevich, Voprosy kosmogonii, Izd. AN SSSR, No. 5 (1956).

BOOK REVIEWS

Studies in the History of Astronomy. No. 1. P. G. Kulikovskii, Ed. Gostekhizdat, Moscow, 1955. 367 pp, 12 rubles, 10 kopecks. No. 2, Moscow, 1956. 420 pp, 12 rubles, 80 kopecks.

The series of Studies in the History of Astronomy which began to appear in 1955 is of varied content. It includes articles of a historical nature, sketches concerning the history of Russian astronomical observatories and the development of various kinds of astronomical work in Russia and the USSR, memoirs, biographies of the outstanding figures in astronomy, the correspondence of Russian scientists, reference materials in the history of astronomy and annotated bibliographies.

The first issue contains ten varied articles; four articles are devoted to the eminent Russian astronomer I. M. Simonov, two articles concern V. K. Tseraskii, there are letters from A. M. Zhdanov to V. V. Vitkovskii and a bibliography of the literature on the history of astronomy which appeared in the USSR during 1953 and 1954.

The first article, by V. L. Chenakel, describes documents relating to two plans for the construction of an astronomical observatory in St. Petersburg to replace the academic observatory which had been burnt down in 1747 and which had been located in the tower of an art museum. These plans were formulated by Professor Grishov, who had been appointed in 1751, and contain very advanced ideas regarding the setting-up of instruments not in a high tower but in a low building on solid foundations insulated from the floor and walls of the building. The plans and building sites on Vasil'evskii Island failed to obtain approval and the academic observatory remained at its previous location for 75 years. This article is of interest for the history of astronomical observatories in Russia and calls for only one minor comment: The author twice (on pages 22 and 45) uses the word "liuk" in the sense of the shutter which covered the meridian slit in the observatory wall, whereas strictly speaking this word designates the opening itself, which was covered by a lid or shutter.

P. V. Slavenas' article "Astronomy in the Higher School of Lithuania during the 16th to 19th centuries" is based on extensive printed and manuscript material and contains interesting information regarding the "Higher School" of Lithuania and the Vilna astronomical observatory, which existed more than a hundred years before the fire of 1876 and which during its last years was advanced in the use of photography for studies of the sun and moon.

The following article by G. D. Dzhalialov, "The Difference between the Guragoni Tables and Other Similar Tables," contains a list of various astronomical tables of the medieval Orient among which the author strangely includes the tables of the ancient Greek scholars Savin of Alexandria, Hipparchus and Ptolemy, which is justified neither chronologically nor geographically. The author compares the successive titles of chapters and sections of three different sieges compiled in Samarkand, Maragha and Delhi. Without analyzing the contents of the Samarkand tables, which are called the Guragoni Tables, the author uses a purely formal criterion – the identity of designations of divisions and sections – and follows the opinions of other authors to distinguish the tables which he regards as the independent work of Samarkand scholars. These are tables of the trigonometric functions as well as tables of the sun, moon and five planets and "tables compiled principally on the basis of the planetary tables". The last part of this sentence is entirely incomprehensible and the conclusions of the author are unconvincing as a whole. A solution of the problem required not a comparison of the titles of the sections but a more profound analysis of the contents of the tables, of the theory on which they were based and of the constants used for the planetary orbits. This is the only way of determining the amount of borrowing from other sources and the independent contribution to planetary theory and ephemeris astronomy of that period.

Similar criticisms must be made of another article by the same author entitled "On the Compilation of the Planetary Tables of the Samarkand Observatory". This article discusses principally the famous Ulug Beg observatory in Samarkand rather than the tables. There are included somewhat primitive diagrams and a description of the principal instrument of that observatory, which was a large sextant. It may be mentioned that for a single degree of length 70.2 cm the radius of this instrument was 40.22 m rather than 40.42 m as stated on p. 110 and certainly not 44 m as stated on p. 12 of the article by M. K. Venttsel in the second issue of the Studies. The author quotes Abd ar-Razak to the effect that the Samarkand observatory contained an "object divided into degrees, minutes, seconds and tenths", which the author identifies as a large sextant. But from the state of technology of the time it is clear that no instrument could have been in existence with such small divisions and also that such divisions would have been entirely useless since viewing with the naked eye could not be more accurate than to within a few minutes. It is evident that the manuscript of Abd ar-Razak was falsely interpreted or that a different meaning must be given to his words.

The Hungarian scholar D'erd' Nador in his article "Kepler's World View" shows convincingly that, contrary to the opinion of certain bourgeois historians of science, Kepler was no Pythagorean mystic but a bold materialist who sought the objective nature of the laws of nature.

Iu. G. Perel's article on V. K. Vishnevskii contains some errors. Thus, the burning of the academic observatory in St. Petersburg occurred in 1747 rather than in 1743 (p. 134). In speaking (p. 137) of Vishnevskii's observations of the comet of 1811, at a time when it had been lost by all other observers, the author attributes this to Vishnevskii's skill and does not mention the principal fact that Vishnevskii possessed unusually sharp vision. It is also stated that these observations were made under the unfavorable conditions of an expedition rather than in the St. Petersburg observatory. But this is precisely what made it possible to observe the comet until August 1812, as the white nights of the summer prevented viewing of the comet in St. Petersburg, which Vishnevskii observed in the dark sky of Novocherkassk. There is an exaggerated statement that the observations of eclipses of Jupiter's satellites required equipment everywhere in the contemporary observatory (p. 138). For this purpose a three-inch portable telescope was sufficient. On p. 141 there is an obvious error or misprint in the comparison of the accuracy of determining geographical latitudes — 5", and of longitudes — 2", We know that at the time longitudes were determined much less exactly than latitudes: instead of 2" of arc 2 seconds of time should be mentioned, which is 15 times greater. Otherwise the article provides a rich fund of information.

Some thorough and valuable articles present the history of observatories and the development of astronomical investigations in a number of universities. D. V. Piaskovskii writes of the development of astronomy at the Kiev University. S. N. Korytnikov informs us of the beginnings of astronomy in Kazan and in the second issue of the Studies gives an interesting account of the early period of astronomy at Moscow University. Ia. Iu. Korpun and V. P. Tsesevich in a biography of A. K. Kononovich in the second issue give a detailed history of the Odessa observatory.

In connection with the first of these articles we must note that in addition to courses and the textbooks mentioned on p. 175 Khandrikov wrote the following books: The General Theory of Perturbations, A Course in Spherical Astronomy, An Outline of Theoretical Astronomy and A Comparison of the Methods Used by Gauss, Bessel and Hansen to Calculate Solar Eclipses. It is unfortunate that the illustrative material in Piaskovskii's article is so poor; the slanted position of the tower in the photographs on pp. 174 and 186 and the interfering telegraph pole as well as the poor view of the principal building of the Kiev observatory (p. 167) make an unfavorable impression.

The article on A. K. Kononovich contains much documentary material, especially a report of N. A. Umov and F. N. Shvedov on his doctoral dissertation "A Photometric Study of Mars, Jupiter and Saturn", showing that Kononovich ignored the different colors of the photometrized planets and comparison stars Altair and Vega, took into account only very approximately the atmospheric extinction at low altitudes of the observed planets, made an insufficiently rigorous investigation of various instrumental errors and committed some mathematical errors in calculating integrals. This is hard to reconcile with the opponents' conclusion (p. 333): "The book is a clear proof of complete acquaintance with the literature of the subject and with methods of measurement and calculation and is also a proof of the diligence of the author". It is surprising that Kononovich received his doctorate after this report.

In the same article the following sentence is striking (p. 335): "A study of the methods of astronomy was at that time required of all students in the mathematical division of the physico-mathematical faculty". This was in the eighties of the past century. It is regrettable that this well-justified requirement does not prevail at the present time.

Returning to the first issue, we consider the article by V. L. Chenakal on the "Large Fixed Telescope of Lomonosov with a Siderostat". The first sentence, which ends with ". . . . for the reflection of light of heavenly bodies moving in a constant direction" obviously requires the following rearrangement:" for the reflection in a constant direction of light from moving heavenly bodies". Of course it would be interesting to learn the construction of Lomonosov's siderostat. Recently discovered documents enable us to determine the time when Lomonosov was occupied with this matter, the name of the technician who made the flat mirror for the siderostat and the size of the mirror. But unfortunately this information is of least interest and does not justify the author's statement (p. 221) that it ". . . gives a complete picture of the work of the great scientist on this remarkable invention". The design of the siderostat and the clock mechanism remain entirely unexplained, although these are the most important facts about the instrument; the size of the mirror is a secondary unimportant detail. The article does not give entirely correct information about the invention of the siderostat. On page 208 it is stated that priority in the introduction of the siderostat into practical astronomy is attributed by western European historians of science to Gauss, Repsold and Radau. But the design of a heliostat, that is, a siderostat for solar observations, was described in 1620 by the Dutch scientist s'Gravesand Gauss, as we know, was the inventor of a heliostat for an entirely different purpose, Repsold was a German specialist in mechanics who among many other astronomical instruments made a few siderostats and Radau was a French astronomer who in 1884 described the long-known principle of the siderostat and a new design of this instrument proposed by Foucault, Silberman and Littrow. Thus none of the authors mentioned by Chenakal has a claim to the assigned role, since s'Gravesand must be regarded as the first inventor of the siderostat, and the inventor of the most widely used form was the French physicist Foucault. At the present time almost none of these instruments is used in astronomy; they have been replaced by the much more perfect coelostat invented by August in 1839 and introduced into practical astronomy at the beginning of the 20th century by Lippmann and Hartmann.

M. E. Nabokov in a detailed article entitled "An Outline of the History of Instruction in Astronomy in the Secondary Schools of Russia and the USSR" shows that the curriculum of prerevolutionary secondary schools included one to two hours per week for "cosmography" in the senior year (this was the name given at that time to elementary general astronomy). Unfortunately this subject has often been greatly curtailed in present secondary school curricula and has been retained only with great effort in the curricula of the RSFSR. A large bibliography is attached to the article.

The last section of the first issue contains documents and other material on the eminent Russian astronomers I. M. Simonov and D. M. Prevoshchikov,including Simonov's autobiography and Perevoshchikov's letters to him. There are two interesting articles on V. K. Tseraskii. In the first of these Iu. G. Perel' reports Tseraskii's utterances on general political subjects taken from handwritten notes. This great scientist is thus shown as a progressive individual who had experienced the subjugation of his native Poland by Czarist Russia. In the second article F. Ia. Zotov reminisces regarding the last and least known years of Tseraskii. Unfortunately these recollections are quite scanty since the author met Tseraskii only seldom and then by chance.

The first volume ends with letters of A. M. Zhdanov to V. V. Vitkovskii with notes principally of a biographical character. There are some omissions and inaccuracies in these notes: On p. 354 there is no mention of Vitkovskii's well known books "Topography", the fourth edition of which was issued in 1940, and "Experiences", which is autobiographical and therefore especially important for the book under review. On p. 355 concerning Bolin there should have been some mention of his method of calculating the movements of the small planets. On p. 356 no mention was made of the principal work of Charlier, his statistical stellar investigations. Concerning Harzer it is not stated that after working at Pulkovo and Gotha he was director at the Kiev observatory from 1897 to 1926. On p. 357 it is not at all clear what measurements of the large planets were obtained by Struve using the 76-cm Pulkovo refractor (in actuality he measured the positions of the planetary satellites and developed a theory of the movements of Saturn's satellites). Regarding Zinger it is not stated that he was the author of a well-known and widely-used method of determining time and in connection with Newcomb (p. 358) his most important work is omitted – the determination of the elements and the compilation of tables of motion of the four inner planets as well as the determination of astronomical constants. Finally, certain names encountered in the letters are not explained. Thus on p. 346 mention is made of Ederin and his "method which will be used by Russian geodesists"; it should have been explained that Ederin, a Swedish geodesist, invented a method of measuring bases by means of wires, which is what Zhdanov's letter refers to.

The most important article in the second issue, "A Brief Outline of the History of Practical Astronomy in Russia and the USSR", was written by M. K. Venttsel and fills 127 pages. At the beginning of the article there is a discussion of the instruments of antiquity and of the Renaissance, which is outside the scope of the article. The very first sentence is objectionable: it is stated that the instruments of ancient astronomers were "so primitive and inaccurate, that it is almost improper to speak of their use for determining geographical locations". But Ptolemy's "Geography" contains a list of almost 8000 geographical locations, since it was well known at the time that the drawing of geographical maps requires astronomical determinations of positions. The inaccuracy of the instruments and the primitiveness of the methods resulted only in inaccurate geographical coordinates, especially longitudes, but certainly did not invalidate astronomical observations for this purpose. On p. 13 the island where Tycho Brahe built his famous observatory should be transcribed in Russian as "Ven" and not "Khven", in accordance with Danish pronunciation (English — Hven). The statement that for determination of latitude from the noonday altitude of the sun (p. 21) the longitude of the point must be known is true, but it should have been explained that this longitude need be known only roughly since in the most unfavorable instance (at an equinox) an error of 15° in longitude produces an error of only 1' in latitude, so that this method could have been used successfully at the time of Ulug Beg. On p. 26 there is an error in the conversion of degrees to time: 1°.4 = = 5.6 min rather than 21 min as printed. On p. 34 the accuracy of the observations of Peter I is given as of the order ± 0.5 instead of ± 0.05' sec of time. The article contains too little of the remarkable period of Russian geographical determinations of the 18th century. The names of the outstanding astronomers and travelers P. V. Inokhodtsev, I. Islen'ev, L. Kraft, Khr. Eiler, the unfortunate Lovits and the calculator Leksel' are not even mentioned. It would be very interesting to have a detailed account of the method of determining astronomical points at that time. The bibliography, which is quite full in general, does not mention the beautiful sketches on the development of astronomy in Russia before the founding of the Pulkovo observatory, which were written by V. Ia. Struve and printed in the "Description de l'Observatoire de Poulkova" and in the Journal of the Russian Geographical Society for 1849. P. 55 gives an incomplete and loose idea of the conception of an absolute star catalogue; the principal feature is not mentioned, which is an independent determination of the equinox and equatorial point. The account of the history of the Moscow Surveying Institute abounds in minor details while the principal work of Shveitser on the investigation of the Moscow attraction is mentioned only once (p. 62). The different methods of determining latitude and time which have been developed mainly by Russian and Soviet scientists are described in detail and are perhaps the principal merit of the entire article. It should have been mentioned, in speaking of the Astronomical Yearbook of the USSR (p. 104), that at present it is compiled at the Institute of Theoretical Astronomy of the USSR Academy of Sciences entirely independently of any borrowing from any foreign ephemeris.

In the article by V. L. Chenakal on the "Astronomical Observatory of the St. Petersburg Academy of Sciences at the End of the 1730's" the plan of the various stories of the observatory is given together with a list of the instruments according to original source material. This shows once more how well and completely the observatory was equipped.

The next article by the same author, entitled "The St. Petersburg Meridian" tells of the establishment in 1774 of the St. Petersburg meridian marker which, in the words of the author, "was of great practical importance in the history of science and in general for the Russian nation". The article states that three different purposes were served: the fixing of the St. Petersburg meridian, measurement in degrees along the meridian and checking of magnetic instruments. But fixing of the meridian by a special marker in connection with the computing of longtitudes in Russia from St. Petersburg is not at all required, because longitudes associated with the determination of exact time are always computed from an astronomical observatory and the St. Petersburg Observatory fixed the meridian to a sufficient extent. The concept of a personal equation was nonexistent at the time and no need could have arisen for a fixed point to determine the personal equation. The second possible purpose is discussed in a quoted letter of Delisle, but since no degree measurements were made along the St. Petersburg meridian nor along the corresponding parallel the purpose is not proved. Only on the eighth page of the article does it become clear why the meridian marker had to be fixed; from a declaration of the military geodesist D'iakov it appears that it was a matter of checking magnetic instruments and that for this purpose not only a fixed point of the meridian was needed but also a fixed direction of the meridian line or at least known azimuths of observable points within the limits of visibility from a meridian. The meridional line on the top slab of the column which was set up could not serve for this purpose, because magnetic compasses had a sighting device for aiming at distant points and the short meridian line on the slab was insufficient for an exact orientation of the magnetic needle, and all the more so because a sundial was later set up on the same slab. It is thus not surprising that we have no record of the utilization of this mark by Russian geodesists. We must suspect that the mark was unsuited for the purpose, although the author in the last sentence of his article speaks without proof of the practical use of the mark by Russian astronomers, geodesists and physicists. In this article no mention is made of any exact astronomical fixing of the mark; in this respect also the final sentence is not confirmed.

Further articles by S. N. Korytnikov concerning the early history of astronomy at Moscow University and the causes for the departure of D. M. Perevoshchikov from the university as well as letters of Perevoshchikov contain extensive and interesting source material.

Recollections of the first year of the late B. P. Ostashchenko-Kudriavtsev at the Pulkovo Observatory contain almost no information of a scientific character, but are interesting from an everyday point of view as a description of life at the "astronomical capital of the world" during the first years under O. A. Baklund as director.

At the end of the second issue there is an index of names mentioned in the text, which is unfortunately absent from the first issue.

The above brief account of the articles shows the diversity of the contents of the first two issues of the Studies. Most of the articles are of great value for the history of the development of astronomy in our country. These are the first printed accounts of little known or entirely unknown periods of the past, lives of outstanding astronomical scientists and histories of astronomical observatories. It was no easy task to collect all of this material and we can congratulate the editorial board headed by N. G. Kulikovskii, the authors and the Gostekhizdat on a great and useful achievement. We hope that there will be further issues of the Studies in the History of Astronomy.

Received March 19, 1957

A. A. Mikhailov

Iu. A. Riabov, The Motion of the Heavenly Bodies. Gostekhizdat, Moscow, 1956. 159 pp. 2 rubles, 65 kopecks.

Popular scientific literature is very important in our country. Many books are published but they are far from illuminating uniformly the various branches of our science of astronomy.

Let us take the sun for example. On this subject sometimes, although rarely, popular scientific books are written by great scientists who are themselves conducting serious investigations in the field, or by authors who although they themselves are not directly occupied with the subject understand it thoroughly. Sometimes however, books are written by authors whose 'knowledge is not sufficient to give them the right to do so. Unfortunately books in this third group are printed. In this way much popular scientific literature is issued regarding the structure of the solar system, the structure of the universe and some other branches of astronomy. But some branches are neglected, among which is celestial mechanics. This is quite natural: not every author who has read several books or articles on celestial mechanics will undertake the task of writing a popular book on the subject. Therefore the appearance of the popular book by Iu. A. Riabov, "The Motion of the Heavenly Bodies", must be welcomed. This book explains in popular form many very interesting questions concerning the motion of the heavenly bodies. This is especially important at the present in connection with the construction of artificial satellites and the problems of interplanetary communication, in which very many are interested and on which special institutions and individual scientists are working. The book gives a reasonable account of the law of universal gravitation and an idea of perturbed motion. There is a necessarily elementary account of the theory of the motion of the heavenly bodies and at the end a discussion of the nature of gravitation. On the whole the book is interesting and in our opinion necessary. It can be read with profit by anyone who wishes to become acquainted with the problems of the motions of heavenly bodies. However, the book contains defects to which we shall here devote most of our attention.

The general defect of the entire book is the unclear exposition of the various topics. This may be explained by the fact that the book is the first popular scientific work of the author and that it appeared to him that all were acquainted with celestial mechanics as he himself. He therefore devoted himself mainly to the subject matter and gave insufficient attention to finding methods of explaining the material properly. We shall give a few examples.

On p. 10, in speaking of ancient ideas of the motions of the sun, moon, planets and stars, the author states: "Observations showed that the paths of all the planets among the stars are close to the ecliptic. The planets Mercury and Venus are always comparatively far from the sun, moving away from it in sometimes one and sometimes in the opposite direction, Venus by about 46° and Mercury on the average by 23°. The period of the visible oscillations of Mercury about the sun is 116 days and of Venus one year and 217 days. (It is clear that the author is here talking of the synodic periods of motion of the planets. — K. K.) The apparent motions of Mars, Jupiter and Saturn are not so directly related to the sun. These planets in their motions among the stars describe complete circles with different velocities: Mars in approximately 1.5 years, Jupiter in 12 years, Saturn in 29.5 years". These are sidereal periods, that is, periods of motion of Mars, Jupiter and Saturn around the sun which were not known to the ancients. They assumed the earth to be fixed and they knew only the periods of apparent motions of these planets, that is the synodic periods. The author should thus have written: "Mars in 780 days, Jupiter in 399 days and Saturn in 378 days".

On p. 10 it is stated that "the sun and moon move irregularly and the planets especially described complicated loops". It should have been stated that complicated loops are described only by the outer planets; nothing is said of this.

On p. 11 it is stated that "during the time in which planet P and the center of the epicycle O are moving in the same direction an observer on the earth E sees straight-line motion. When the planet moves between the center of the epicycle and the earth the displacements of the planet and the center of the epicycle "are subtracted" from each other, and from the earth the apparent motion of the planet is retrograde." It must be remembered that retrograde motion of a planet is not always possible but only when the linear velocity of the planet in its epicycle is greater than the linear velocity of the center of this epicycle in the deferent. This should have been stated.

On p. 24 the author states that "Kepler computed the distance between Mars and the sun at 28 points of its orbit and attempted to find a circle which would pass through all of these points". It should have been shown or at least stated to what position of Mars with respect to the earth and sun these points correspond. The discussion by the author indicates only 11 points rather than 28. He speaks of observations of Mars in opposition, which numbered 11 for Kepler.

Pages 86-88. The discovery of Neptune could have been described much more colorfully. It was also necessary to describe the appearanceof Halley's comet which was calculated on the basis of universal gravitation by A. Clairaut. This calculation was performed almost one hundred years before the discovery of Neptune and for the proof of universal gravitation was no less important than the discovery of Neptune. On p. 112 the author states that the forces acting on the earth from the moon and sun are applied to the axis or rotation of the earth. This is incorrect since forces can be applied only to a material body, that is to the earth, and not to an imaginary axis.

On p. 115 the author writes: " The change of length of the day is related to the tides. As we know, the level of the water in the oceans during a day does not remain constant but changes regularly.

For about 6 hours it increases and reaches its maximum (high tide); during the following six hours the water level drops and reaches its low point (low tide). Then a new high tide begins. Each high tide or low tide follows the preceding one on the average by 12 hours and 25 minutes, so that during a day there are two high tides and two low tides repeated regularly at every point of the globe." The author's statement is incorrect: during a day there can be two high tides and one low tide or one high tide and two low tides.

The author's discussion of the increase of the length of the day as a result of tidal friction, on pp. 117-118, could hardly be understood by the reader.

In reporting the fact that as a result of slowing down of the earth's rotation (due to tidal friction) a day of the earth increases by about 0.001 sec in 100 years, the authors writes as follows: "Let one clock follow the rotation of the earth exactly, that is, gradually slowing down by 0.001 sec in 100 years or by 0.001/36525 sec in a day, while another clock is constant. The retardation of the first clock with respect to the second is computed in exactly the same way as we determine the path traversed in uniformly accelerated motion, that is, by the formula $\Delta T = {}^1/_2\, \epsilon\, t^2$, where ϵ is the deceleration of the clock in days, $\underline{t}$ is the time in days. In 100 years, that is in about 36525 days, the retardation of the first clock is about ${}^1/_2 \cdot 0.001 \cdot 36525 = 18$ sec". A more detailed explanation should have been given.

Certain parts of the book are difficult to understand. In Section 20 entitled "What is gravitation ? ", the author attempted to give a popular explanation of both the special theory of relativity and the general theory of relativity. He states that "unfortunately both the special and the general theory of relativity are very difficult to explain in a popular manner". This is true, but when an author undertakes a description it should be done in such a way that the fundamental idea will be comprehensible. This rule was not always observed by the author. For example, on p. 140 he writes that "according to Einstein's theory not only ordinary material bodies but rays of light are attracted by material bodies and are therefore deflected from their straight paths past heavenly bodies, although very weakly; light is not propagated in exactly a straight line. This at first seems strange and contradictory of our daily experience on earth. Nobody had previously observed this. But Einstein was found to be right. During solar eclipses it is possible to observe the stars very close to the edge of the sun. Exact measurements of the positions of these stars during eclipses show that they are not always seen in their usual places but are shifted somewhat in the direction of the sun". This is all that is written of the curvature of light in a gravitational field. Even a person familar with the fundamentals of astronomy would not understand how the curvature of rays is determined by the gravitational field of the sun.

The author should have stated that to settle the question of the curvature of rays from the stars near the sun a photograph is taken. Since the photographing occurs during a total eclipse the photographic plate shows the brighest stars in addition to the solar corona. But photographs of the sun during a total eclipse do not settle the matter. It is necessary that the portion of the sky in which the sun was located during the eclipse be photographed again after a half year when the sun is at a diametrically opposite point of the ecliptic. At that time rays reaching us from the stars do not pass through the gravitational field of the sun.

By comparing these plates the different positions of the same stars make it possible to determine the displacement of these stars during an eclipse.

The book contains obvious misprints. On p. 45 the author writes: "Thus the sun and planets attract each other like material points for two reasons: 1) the distance between them is small compared with their sizes; 2) they are nearly spheres". In 1) he should have written "The distance between the bodies is great".

On p. 138 we find: "This law states that in the absence of forces a body moves uniformly or in a straight line". This should have been: "uniformly and in a straight line".

There are a few other unimportant shortcomings. All of these should be removed in later editions.

Received February 24, 1957

K. A. Kulikov

CHRONICLE

AWARDING OF THE 1957 LENIN PRIZE TO B. K. IOANNISIANI FOR THE DESIGNING OF ASTRONOMICAL INSTRUMENTS

A. N. Dadaev

On April 22, 1957, the 87th aniversary of the birth of V. I. Lenin, the Committee on the Lenin Prize for Science and Technology of the USSR Council of Ministers awarded the Lenin Prize for the designing of new astronomical instruments to Bagrat Konstantinovich Ioannisiani, chief designer of the S. I. Vavilov State Optical Institute.

B. K. Ioannisiani was born in Erevan in 1911 and in 1929 was graduated from a 10-year school in Leningrad. His working life began at that time. He entered the "Red Dawn" factory as a draftsman and without interrupting his work completed courses in designing. He then worked at the GOMZ (State Optical Equipment Plant). At that time he became acquainted with N. G. Ponomarev, who became interested in the young designer and initiated him into the technique of producing astronomical instruments. During the war years Ioannisiani worked as a designing engineer in Kazan. In 1945 he became chief designer of the GOI (State Optical Institute). Here in collaboration with D. D. Maksutov he began on introducing an increasing number of meniscal systems into astronomy. He is responsible for the large number of new and original designs of astronomical instruments which were produced by the Soviet Optico-mechanical industry in the last few years and which marked the genuine flourishing of our domestic telescope designing.

His first success was the designing of an ASI-1 nebular spectrograph for the Siemens and Biurakan observatories (1949). The distance from the slit to the prism of the spectrograph is 50 meters; this instrument is thus the largest of its kind. The camera has a powerful meniscal system with 1:1 relative aperture. Soon afterward he carried out Maksutov's idea for the outstanding ASI-2 meniscal telescope of 500 mm diameter (1950). This telescope is at the Alma-Ata Observatory, to which the designer himself traveled to install and test the instrument. This telescope is designed for photographing at the principal focus with a 1:2.4 relative aperture (f = 120 cm). V. G. Fesenkov and D. A. Rozhkovskii used it to produce an atlas of gaseous nebulosities and made especially careful investigation of filamentary nebulae for the purpose of studying possible processes of star formation.

There followed a series of original but smaller instruments: a lens-mirror camera (ASI-4), a reflecting telescope with a slitless quartz spectrograph (ASI-5) and a slitless meniscal diffractive lens spectrograph (AS-31). These instruments are described in the literature. We must also mention the designing of the stellar electrophotometer, of which a few copies were produced by the shops of the Physical Institute of the Leningrad State University, and also experimental telescopes produced under his direction at the Main Astronomical Observatory in Pulkovo.

An outstanding achievment of Ioannisiani as a designer is the AS-32 meniscal tesescope for the Abastuman Astrophysical Observatory. We shall give some details of this instrument which are not yet in print. The entrance aperture of the telescope (the light diameter of the lens) is 700 mm and the diameter of the principal mirror is 975 mm. This telescope is designed for direct photographing at the principal focusing with a 1:3 relative aperture (f = 210 cm) and with a 5°.5 diameter of the effective field ($18 \cdot 18$ cm^2 plates), and also for photographing with an objective prism of 720 mm diameter and about 8° angle of refraction. This is the largest objective prism in the world. Two auxiliary lens compensators permit observations in a Maismith focus with 1:15 relative aperture (f = 1050 cm) and 40' and 20' field diameter. For observations in the secondary focus a diffraction spectrograph can be used with two rotatable lattices providing a resolving power of 24 and 80 A/mm. This spectrograph possesses high thermal stability. The expanding of the spectrum is automatic.

The instrument is equatorially mounted in a yoke. The telescope is guided automatically. Pointing at an object with given coordinates is accomplished by remote control from a separate panel. The right ascension and declination are given directly and the hour angle is calculated automatically by a special clock mechanism. The control panel contains adjusting dials and conversion dials (by right ascension and declination) as well as hour angle and stellar time dials. Adjustment of the instrument requires about three minutes;the shifting rate is 60° per minute. There is also automatic timing and cessation of exposure in photographing from 5 seconds to four hours.

Two meniscal guides with a focal length of 250 cm are used which permit inclination at the angle of deflection of the prism. In the secondary focus guiding is accomplished by means of a Ritchey cassette.

Preliminary testing (after installation of the telescope in the tower and its adjustment) during October 1956 in Abastuman with the participation of Ioannisiani showed the high optical qualtities and design data of the telescope. Soviet and foreign astronomers at the plenum of the Commission on Solar Investigations in November 1956 at Tbilis who visited the Abastuman Observatory gave high praise to the instrument.

In the construction of this largest meniscal telescope the following participated: designers B. F. German, M. K. Iarotskaia (who developed individual units), B. Ia. Karasik (who designed the control panel), master opticians M. A. Stepanov and L. V. Simonenko, technicians K. K. Shulepov and S. I. Lebedev and others.

At the present time Ioannisiani is working on the construction of an even more powerful and improved telescope of 2.6 meter diameter for the Crimean Astrophysical Observatory. All Soviet astronomers wish him success.

Received April 26, 1957

PLENUM OF THE COMMISSION ON THE HISTORY OF ASTRONOMY

Iu. G. Perel'

On February 6, 1957 during the annual plenum of the Astronomical Council of the USSR Academy of Sciences at Pulkovo a regular plenum of the Commission on the History of Astronomy was held. The following were present: Prof. B. A. Vorontsov-Vel'iaminov, Corresponding Member of the USSR Academy of Sciences M. S. Zverev, E. L. Krinov, Prof. B. V. Kukarkin, P. G. Kulikovskii, G. D. Memedbeili, Prof. D. Ia. Martynov, Prof. O. A. Mel'nikov, Corr. Member A. A. Mikhailov, A. A. Nemiro, Prof. K. F. Ogorodnikov, B. A. Orlov, Corr. Member P. P. Parenago, Iu. G. Perel', Prof. P. V. Slavenas, Corr. Member M. F. Subbotin, V. L. Chenakal, Prof. V. P. Shcheglov, Prof. N. S. Iakhontova and others.

The Chairman of the Commission, P. G. Kulikovskii, reported on its work in 1956 and on various writings in the history of astronomy by members of the Commission and by other Soviet investigators during 1956 or which were printed during that year. An important event was the printing of "Outlines of the History of Russian Astronomy" by Prof. B. A. Vorontsov-Vel'iaminov. The office of the Commission took an active part in preparing this work for publication and in reviewing and editing it.

It was noted that unfortunately many members of the Commission did not complete their assigned work in the history of astronomy. This resulted from an excessive burden of their principal work and from the fact that as a rule historical studies are not included in the research plans of observatories and institutes, being performed as extra work.

The principal task of the office of the Commission was the publication of "Studies in the History of Astronomy". The first issue of these "Studies" appeared at the beginning of 1956; the second issue appeared at the beginning of 1957. Their contents reflect the varied interests of Soviet students of the history of astronomy. The second issue contains articles by foreign scholars — Prof. E. Rybok of Poland and D. Nador of Hungary. Kulikovskii reported on the plan for the third issue, which will appear during 1957. The fourth issue, for which some material is already available, is planned for publication during 1958 for the 10th Congress of the International Astronomical Union. The Commission established relations with historians of astronomy of China and other popular democratic countries. In August 1956 Kulikovskii reported at the Astronomical Conference in Budapest on Soviet work in the history of astronomy. The Commission arranged the participation of Soviet astronomers at the International Congress on the History of Science held in Florence, Italy during September 1956; four reports were presented. The Commission also arranged, although inadequately, for the reviewing of manuscripts in the history of astronomy. These included a chronology of astronomical discoveries for a supplementary volume of the 2nd edition of the Great Soviet Encyclopedia.

The reported year failed to see the publication, planned by the preceding plenum, of descriptions of the scientific archives of astronomical institutions. No progress was made on the compilation and publication of a biobibliographical dictionary of Russian and Soviet astronomers. The office of the Commission will request the cooperation of the plenum of the Astrophysical Council in the first matter and will make suggestions regarding the compilation of the dictionary.

A lively discussion ensued. Some speakers criticized the decision of the Glavknigotorg to publish the second issue of the "Studies in the History of Astronomy" in only 1500 copies and showed that such a unique work should be printed in not less than 3000 copies. The meeting voted to appeal to the plenum of the Astronomical Council for support in the matter of increasing the printing of the "Studies in the History of Astronomy". V. P. Shcheglov commented on the need for directing the development of the history of astronomy in the Central Asian republics, which had a glorious astronomical past. Many speakers emphasized the need for collecting information on the existence of source material in the history of astronomy in archives of Soviet astronomical institutions and in government archives and for the publication of a description of such archives in the Studies.

There was a long discussion of the advisability of publishing letters of A. A. Belopol'skii to V. K. Tseraskii, which are in the archives of the GAISh (Shternberg State Astronomical Institute) and which contain very subjective opinions concerning the conduct of F. A. Bredikhin as director of the Pulkovo observatory. It was finally decided that it would be inadvisable to publish these letters at the present time.

Finally, there was a discussion concerning further work on the compiling of the biobibliographical dictionary of Russian and Soviet astronomers. The Scientific Secretary of the Commission, Iu. G. Perel', reported on suggestions of the office of the Commission. A. A. Mikhailov commented on the need for including an indication of the principal scientific works in the biobibliography of astronomers. The meeting accepted the suggestion that the work of compilation should be divided among astronomical institutions. The office of the Commission proposed that a list of astronomers to be included in the dictionary be sent to observatories and institutes for their comments. Following this the Commission would be able to formulate a final plan for distribution of the work, which should be finished during 1957 and published during 1958.

This lively meeting will undoubtedly further the development of the history of astronomy.

Received March 30, 1957

REPORTING AND COORDINATING PLENUM OF THE ASTRONOMICAL COUNCIL OF THE USSR ACADEMY OF SCIENCES

V. S. Safronov

On February 5-7, 1957 at the Main Astronomical Observatory of the USSR Academy of Sciences at Pulkovo there was held a plenum of the Astronomical Council of the USSR Academy of Sciences for a report of its activities during 1956 and for coordination of the plans of scientific investigations of the astronomical institutions of the USSR in 1957. The plenum was attended by members of the Astronomical Council, heads of commissions of the council and the directors of the astronomical institutions of the USSR Academy of Sciences and of the Academies of Sciences of the united republics and of the Ministry of Higher education.

A report on the activity of the Astronomical council was made by the chairman of the Council, A. A. Mikhailiv, Corresponding Member of the USSR Academy of Sciences. He remarked that most of the work on the coordination of plans is being carried out by the commissions. A plenum of each commission discusses the main lines of investigations, programs of complicated studies and the most important results. The commissions consider plans for scientific investigations by the observatories and make recommendations. In 1956 the most active were the Commission on Solar Investigation and the Astrometrical Commission. Much work was done by these commissions in preparation for the International Geophysical Year, with the participation of the Commission on Comets and Meteors and the Commission on Radio Astronomy. One of the important forms of coordination consists in scientific conferences concerning the most important problems of modern science. September 19-24 1956 in Biurakan there was held a joint session of the Astronomical Council and the Armenian Academy of Sciences on the occasion of the opening of the Biurakan Observatory. At that time a conference was held on nonstationary stars in which fifty astronomers participated including eight foreign specialists from China, Mexico, the U. S. A., France and Yugoslavia.

Several conferences have been held by commissions of the Astronomical Council: February 7, at Pulkovo- a plenum of the Commission on the History of Astronomy; February 10-11 at Leningrad – a conference on the physics of the moon and planets; February 10-12 at Pulkovo – a conference on astronomical instruments and new methods in astronomy; April 5-6 at Moscow – a conference on the influence of solar activity on weather and climate; June 29 to July 1 at Odessa – the twelfth conference on variable stars; August 23-25 at Pulkovo – a conference of participants in the international work of compiling a catalog of faint stars at which foreign scientists were present from the German Democratic Republic, China, Poland, Rumania and Yugoslavia; September 28 to October 3 at Tbilis and the Abastuman Observatory – a plenum of the Commission on Solar Investigations at which in addition to a large number of Soviet investigators there were astronomers from Hungary, China, Poland, Czechoslovakia and Yugoslavia; December 3-4 at Moscow – at the conference of the Commission on the Physics of Planets there were preliminary discussions of observations of Mars; December 30 at Moscow – the conference of the Commission on Cosmogony discussed prospects for the development of investigations in cosmology, etc. The Commission on Comets and Meteors organized two seminars: May 28 to June 1 at Moscow on the astronomy of meteors and June 16-30 at Kazan on radiolocation of meteors.

A. A. Mikhailov commented that scientific ties between Soviet and foreign astronomers are being strengthened. The conferences mentioned above were attended by 22 foreign scientists including 17 astronomers from popular democratic countries. During 1956 the Soviet Union was visited by Dr. and Mrs. Brook from Ireland and

Dr. Moller of the USA. Six astronomers from popular democratic countries visited us for several months to do practical work at our observatories and three are working for advanced degrees (aspirants) at the Main Astronomical Observatory and the Crimean Astrophysical Observatory.

Soviet astronomers, for their part, have taken an active part in international astronomical conferences. Soviet delegations traveled to Liege for a conference on "Molecules and Cosmic Objects", where they presented 12 papers; to Stockholm for a conference on electromagnetic phenomena in cosmic physics, where they presented 7 papers; to Budapest for a conference on variable stars, where 5 papers were presented. Five Soviet astronomers spent 2 to 3 months in China to establish closer contacts and organize joint work; two astronomers spent two weeks in Czechoslovakia to become acquainted with the work of Czech astronomers and for consultations concerning the organization of new studies of the sun and photoelectric effects. The Astronomical Council will continue to strengthen and extend its bonds with foreign astronomers.

A. A. Mikhailov commented on inadequacies in the work of the Astronomical Council and its commissions. There is insufficient coordination of the plans of work of observatories and commissions. There should be greater attention paid to developing new methods of investigation and to the development of extensive theoretical and exploratory work. In consideration of controversial scientific matters we must combat intolerance of criticism as well as unobjective criticism. The debate on the report was participated in by P. P. Parenago (Corr. Member USSR Acad. Sci.), M. S. Zverev, V. P, Tsesevich (Corr. Member Ukr. SSR Acad. Sci). B. A. Vorontsov-Vel'iaminov (Dr. Phys. Math. Sci), S. K. Vsekhsviatskii, A. N. Deich, N. A. Kozyrev, B. V. Kukarkin, A. V. Markov, D. Ia. Martynov, K. F. Ogorodnikov, P. V. Slavenas, V. V. Fedynskii, V. V. Sharonov, I. S. Shklovskii, V. P. Shcheglov and others. The speakers commented on deficiencies in the work of the Astronomical Council and the observatories and made a number of suggestions which were accepted by the plenum.

B. V. Kukarkin reported on "International Cooperation in Astronomy". He commented on the considerable increase of scientific ties between Soviet and foreign astronomers during recent years. Close scientific contacts and cooperation are mutually beneficial and should always be strengthened. Trips to foreign conferences should be combined with scientific work at foreign observatories. Foreign astronomers who visit the USSR should be invited to give lectures and to work at our observatories. It is beneficial to organize joint studies, to print articles by foreign astronomers in our journals and to send some of our articles to foreign journals, and also to exchange scientific information by direct correspondence.

P. G. Kulikovskii reported on "Preparations for the 10th Congress of the International Astronomical Union". He presented a preliminary program of the congress and reported on the steps which must be taken in preparation. The organizing committee of the congress (with Academician V. A. Ambartsumian as chairman), which was appointed by the Presidium of the USSR Academy of Sciences, has already started to function and appeals to all Soviet astronomers for every possible assistance. Much preparatory work must be done to insure a clear and uninterrupted procedure for the congress.

February 6 and 7 the plenum discussed the functioning of the commissions of the Astronomical Council. The following reported on the work of the commissions during 1956 and on plans for 1957: E. R. Mustel', A. G. Masevich, V. V. Fedynskii, O. A. Mel'nikov, B. V. Kukarkin, M. S. Zverev, P. P. Parenago, S. E. Khaikin, P. G. Kulikovskii, D. Ia. Martynov and V. V. Sharonov.

G. F. Sultanov reported on the planned construction of the Azerbaidzhan Astrophysical Observatory.

After the discussion of the work of the commissions resolutions were adopted.

The plenum found that the work of the Astronomical Council in 1956 was satisfactory. The plenum pointed out the importance of steps taken in connection with the participation of astronomical observatories in the International Geophysical Year and with preparations for the 10th International Astronomical Congress. The plenum noted the need for speeding up the organization of production of high-quality domestic photographic plates, the introduction of new methods of investigation and the strengthening of university observatories. The plenum called unjustified the attempts to curtail programs in astronomy in universities and secondary schools.

The plenum passed specific resolutions regarding each commission and called for further strengthening of the ties between the commissions and observatories.

Received April 20, 1957